THE
PROCEEDINGS
OF
THE PHYSICAL SOCIETY

Section B

VOL. 64, PART 2 1 February 1951 No. 374 B

CONTENTS

Price to non-members 10*s*. net, by post 6*d*. extra. Annual subscription: £5 5*s*.
Composite subscription for both Sections A and B: £9 9*s*.

Published by
THE PHYSICAL SOCIETY
1 Lowther Gardens, Prince Consort Road, London S.W.7

PROCEEDINGS OF THE PHYSICAL SOCIETY

The *Proceedings* is now published monthly in two Sections.

ADVISORY BOARD

Papers for publication in the *Proceedings* should be addressed to the Hon. Papers Secretary, Dr. H. H. HOPKINS, at the Office of the Physical Society, 1 Lowther Gardens, Prince Consort Road, London S.W.7. Telephone : KENsington 0048, 0049.

Detailed Instructions to Authors were included in the February 1948 issue of the *Proceedings*; separate copies can be obtained from the Secretary-Editor.

THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

Section B

VOL. 64, PART 2 1 February 1951 No. 374 B

An Improved Beam Divider for Fizeau Interferometers

BY H. BARRELL AND J. S. PRESTON
National Physical Laboratory, Teddington, Middlesex

MS. received 23rd June 1950, and in amended form 28th August 1950

ABSTRACT. A transparent material, presumed to be titanium oxide, is suggested for replacing aluminium or silver as a coating for glass beam dividers in Fizeau interferometers, particularly those adapted for measuring length or displacement in the reflected fringe system. The method of depositing the films, and their optical properties, are described; the coatings are very resistant chemically and mechanically. Evidence produced in photographs and microphotometer records, supplemented by calculations of fringe intensity distributions, shows that the performance of the new material in the interferometer is superior to that of aluminium and of plain glass.

§ 1. INTRODUCTION

In the usual arrangement of the Fizeau or wedge interferometer two plane reflecting surfaces, set nearly parallel to and facing one another, are illuminated with parallel monochromatic light at normal incidence. If both mirrors also partially transmit light, two localized systems of fringes (fringes of constant thickness) may be observed, one in light transmitted through the mirrors and the other in light reflected from the mirrors. It is well known that variations of the reflecting and transmitting properties of the mirrors considerably modify the fringe intensity distribution in both systems, the reflected system being especially sensitive to these variations (Hamy 1906, Holden 1949).

The present paper is concerned with applications of the Fizeau interferometer to precise measurements of length and linear displacement. For these purposes an opaque reflector is most frequently used as one of the interferometer mirrors, so that the fringe observations can only be made in the reflected system. The other mirror then acts as a beam-dividing surface which initially separates the incident light into two portions, one reflected and the other transmitted. The latter gives rise to a series of beams, of progressively diminishing intensity, which emerge from the interferometer after successive reflections at the opaque mirror and partial reflections and transmissions at the beam divider. Tolansky (1948, p. 18) has shown, however, that if the extremely fine fringes characteristic of multiple-beam interference are to be obtained with wedge interferometers, the mirrors, besides having a high reflectivity, must also be placed very close together, and the latter condition is clearly impossible to maintain in practical measure-

ments of length or displacement. Often, also, the opaque reflector necessarily has a low reflectivity and it is then necessary to reduce the reflection factor of the beam-dividing surface to obtain maximum contrast and visibility of the fringes. In this case the amplitudes of successive emerging beams fall off more rapidly and the fringes tend to broaden. In reflected systems this broadening often reveals an asymmetrical intensity distribution across the fringes. The asymmetry is particularly characteristic of beam-dividing surfaces produced by depositing very thin metal films, such as films of aluminium or silver, on glass or quartz bases. A more symmetrical but broader type of fringe is obtained if the beam divider is an untreated optical flat. Asymmetry of fringes in a wedge interferometer clearly cannot arise from interference between two beams only, but must result from an out-of-phase condition of the multiply reflected beams. A simple calculation suggests that even when the effective number of such beams is small, their effect on the intensity distribution—whether by sharpening the fringes or by making them unsymmetrical—may not be unimportant in practical applications.

The experiments to be described indicate that the best results for many length measuring applications are secured by using a glass beam divider coated with a film of transparent material, presumed to be titanium oxide. Evidence of its superiority for the purpose is presented in interferograms obtained from a Fizeau type of gauge interferometer, and is supplemented by curves showing the distribution of density across the fringe images recorded on the original photographic negatives. These curves are shown to be good representations of the fringe intensity distributions calculated from a knowledge of the optical properties of the various interferometer mirrors used.

§ 2. BEAM DIVISION IN THE FIZEAU INTERFEROMETER

A Fizeau type of gauge-measuring interferometer is used at the National Physical Laboratory for routine calibrations of slip or block gauges. Such gauges are well known as providing a convenient and practical form of reference length standard for precision engineering purposes: their measurement in specially designed interferometers of the Fizeau and Michelson types has been discussed by Barrell (1948).

The Fizeau interferometer with glass beam divider, as used for the purpose just mentioned, appears to operate effectively as a two-beam interferometer, but the contrast of the fringes is not a maximum. To attain this, neglecting multiple reflections, the intensity of the beam directly reflected from the dividing surface should be made approximately equal to that of the first beam emerging after two transmissions through this surface and one reflection at the other mirror. Assuming the reflection factor of the latter to be 0·4, a value typical of the lapped steel surfaces of slip gauges in the visible region, the two intensities are equal if the dividing surface has reflection and transmission factors, r and t respectively, satisfying the relation $r = 0{\cdot}4\,t^2$. If the surface also has symmetrical optical properties with negligible absorption, then $r = 1 - t$ and $r = 0{\cdot}23$, $t = 0{\cdot}77$ approximately. Values of r and t for a glass flat differ considerably from these.

Owing to the relatively high absorption of films of aluminium deposited on glass, it appears impossible to achieve a higher value of t than about 0·35 for a film having a reflection factor of about 0·3. Nevertheless, a film with these factors seems to yield the best results for length measuring applications, and although the requirements of thickness and optical quality are rather critical, aluminium

has been employed in preference to the more easily deposited silver because of its greater resistance to abrasion and surface contamination. As already mentioned, aluminized and silvered beam dividers of the quality required for gauge interferometry produce an asymmetrical shape of fringe, but the fringes are, nevertheless, more suitable for visual observation than those from untreated glass.

The characteristic fringe shape yielded by a thin metal film deposited on glass or quartz is undoubtedly associated with the asymmetry of its reflection and absorption factors (Rouard 1938, 1949). Better results in beam division might therefore be secured by utilizing the principle of the non-metallic beam divider which has other well-known applications in optics. In this the glass base is coated with a film of transparent material instead of a metal and use is made of interference in the film itself to give the required reflection and transmission factors. If the material has a higher refractive index than that of the glass base, the reflection factor at a given wavelength first rises with increasing thickness to a maximum and then passes through alternate minima and maxima as the optical path in the film further increases. Provided there is no absorption in the film, the minima all have the same value as that of the uncoated glass, while the maxima all attain a value dependent simply on the refractive indices of the material and the glass base. Simple calculation shows that, with glass of index about 1·5, the maxima have a value of 0·25 if the film material has an index somewhat above 2.

§ 3. CADMIUM OXIDE FILMS

Hammer (1948) has shown that certain evaporated or sputtered metallic oxide films have an absorption factor about one-sixth of that of metallic films. It happened that films of cadmium oxide produced for another purpose (Preston 1950) by sputtering were found to have an index of a little over 2. This, together with the ease with which they could be deposited in a controlled thickness, suggested their use in place of aluminium for beam division in the Fizeau type of gauge interferometer.

A film thickness was chosen which brought the first reflection-maximum just into the blue end of the spectrum, the reflection elsewhere in the visible spectrum then being rather below this maximum, but still considerably greater than for an uncoated glass surface. This seemed more convenient, generally, than using a rather greater thickness to bring the maximum more towards the middle of the visible region, for the first minimum would then seriously encroach on the blue end of the spectrum, and would make the film rather inefficient for use in that region. The results with the interferometer were strikingly better, as regards brightness, contrast, symmetry and sharpness of fringes, than those obtained using aluminium films. Unfortunately, however, the cadmium oxide films proved very susceptible to chemical attack by dust particles, finger-prints and so on. Other materials were then considered, including the oxides and sulphides of zinc and lead. The sulphides in particular have, of course, been used previously for beam division in other optical instruments. It seemed worth while, however, to look for a material even more resistant chemically, and if possible also harder, in view of the constant use to which the interferometer flats were subject. Titanium dioxide seemed to have these properties and so attempts were made to deposit it as a thin film.

§4. TITANIUM OXIDE FILMS

Only methods of coating the glass flats 'in the cold' were envisaged. Cartwright and Turner (1939) refer to the production of semi-reflecting films of TiO_2 either by evaporation or by conversion of hot $TiCl_4$ vapour in air, but no experimental details are given. Other methods have come to the authors' notice in the course of this work (e.g. Banning 1947), but these involve heating the glass to at least 200° C. This is undesirable with optical flats of high quality, so that the method now to be described loses none of its immediate interest.

Of many attempts to sputter titanium as metal or oxide, using a sheet of the metal, none was successful, and a high-vacuum evaporation process was then tried, with success. It follows closely the same lines as for aluminium films. A tungsten wire helix of, say, six turns of 0·7 mm. diameter wire is prepared on a mandrel of about 6 mm. diameter. It is stretched to give the turns a spacing of a few millimetres and mounted in a high vacuum plant exactly as for aluminizing. It is cleaned by heating *in vacuo* in the usual way. A paste is then made with titanium dioxide powder and clear cellulose lacquer diluted tenfold or more with a half-and-half mixture of acetone and amyl acetate. This paste is mixed to the consistency of quite thick cream and painted as evenly as possible over the turns of the tungsten helix. The helix is then warmed gently, by electrical means, to evaporate the solvents, and afterwards brought to a dull red heat for a few minutes to burn off the cellulose; all this is done in air.

Then, in a hard vacuum, the helix is run at a fairly high temperature with a current of 38–40 amperes for perhaps a minute. The white oxide is then found to have sintered into a glassy, black coating on the helix. The glass surfaces to be coated are then chemically cleaned, dried with clean cotton fabric, and mounted in the vacuum chamber at a suitable distance, say 10–15 cm., from the helix. After the usual 'clean-up' at backing pressure, using a high-voltage discharge in the chamber, the high-vacuum evaporation is made at the same temperature as the sintering process. No doubt both these last two stages could be carried out with one pumping of the chamber, if a suitable shutter were used between the helix and the glass flats during the sintering.

Deposition of the film is continued until its colour by transmitted light is a pale brownish grey; if the film becomes a little too thick its colour changes to a clearer very pale blue. A fair amount of gas is often liberated during the deposition, which may be interrupted by switching off the helix; this may be necessary if the pressure rises too high or if it is desired to inspect the progress of the film. Generally, a pressure range of 0·1 to 0·5 micron, as indicated by an ionization gauge, seems satisfactory.

§5. PROPERTIES OF THE TITANIUM OXIDE FILMS

The films thus obtained were found to be very resistant mechanically, and unaffected in a short time, at least, by cold aqua regia or caustic alkalis. The black material left on the tungsten helix was found to scratch glass quite easily, but its exact composition, and that of the films, has not yet been examined. Figure 1 shows the spectral reflection and transmission curves for a glass slip coated with a typical film, the plotted values being uncorrected for the effect of the uncoated surface of the glass. The curves relate to normal incidence. Also, in the same figure are shown corresponding curves for a thin aluminium film

on glass—typical of those prepared for beam division in Fizeau gauge interferometers—showing the asymmetry of reflection which gives rise to the practical disadvantages already mentioned.

It is clear from the figure that the variations, through the spectrum, of the properties of the oxide specimen are of no great importance, and that the oxide specimen has nearly the desired reflectivity. At 5,461 A. the reflectivity of the coated glass slip is about 25 per cent. Thus, after allowance is made for the contribution of the free surface of the glass, and of black paper placed behind it during the reflection measurements, the reflectivity of the actual film is about 22 per cent.

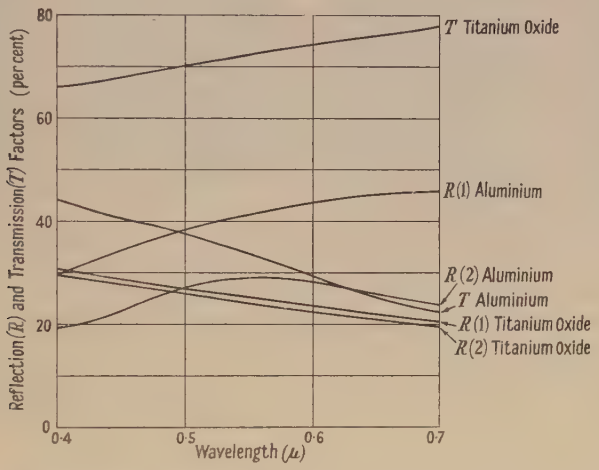


Figure 1. Spectral reflection and transmission curves for films of titanium oxide and aluminium on glass (normal incidence); light incident (1) on coated face of glass specimen, (2) on uncoated face of glass specimen.

Finally, the light lost by absorption in the oxide film is considerably less than for the aluminium film. This may be inferred directly by summing the reflection and transmission factors for the coated specimens and subtracting from unity, correction for the paper backing being unimportant in this comparison.

§ 6. COMPARISONS OF BEAM DIVIDERS

A direct comparison of the fringes produced by using different beam-dividing surfaces in a Fizeau type of gauge interferometer was made in the following manner. A central strip of the beam-dividing surface of a glass optical flat was left untreated and the areas on either side of the central strip were suitably coated, one with an aluminium film and the other with a titanium oxide film. The interferograms obtained in the four cadmium lines are reproduced in Figure 2, the gap between the beam-dividing surface and the lapped steel mirror being about 3 mm. In each interferogram the left-hand strip is aluminized and the right-hand strip is coated with titanium oxide; between these lies the untreated area which is rather wider than was originally intended.

Measurements of the distribution of photographic density across the fringe images were made in a microphotometer employing as detector a barrier-layer photocell associated with a logarithmic potentiometer giving direct readings in density (Hall 1949). Density distribution curves obtained by plotting density

(above fog level) against traverse for each of the twelve fringe groups of Figure 2 are given in Figure 3. Features of symmetry and asymmetry, as well as of contrast and width of the minima, are readily identified. As a matter of interest the dotted line curve for aluminium in red light has been inserted in Figure 3 because it represents a more typical result obtained with an aluminium film of the quality best suited for length measurement. Tests made with a stepped density wedge indicated that the relation between density and log (intensity)

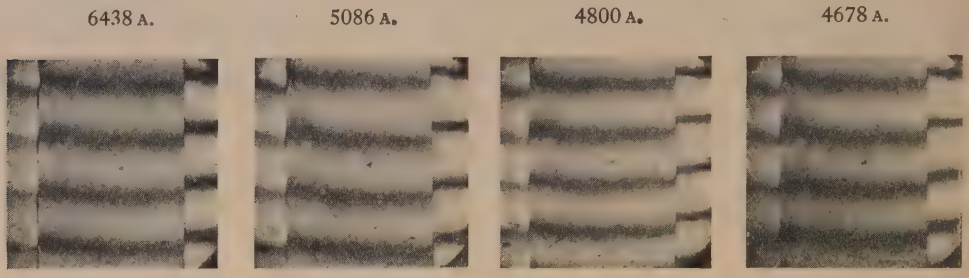


Figure 2. Fizeau interferograms in four cadmium lines; the areas of the glass beam divider to left and right of the central band were coated respectively with aluminium and titanium oxide.

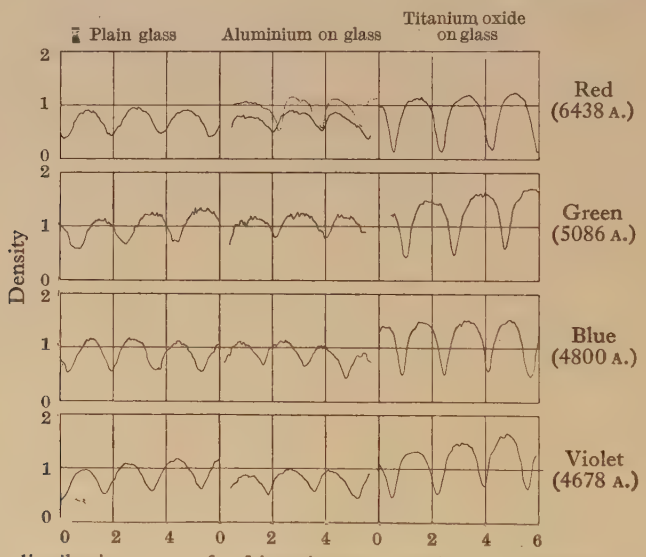


Figure 3. Density distribution curves for fringe images recorded on photographic negatives.

was reasonably linear for densities ranging from 0·25 up to the highest value recorded on the photographic negative. The diagrams in Figure 3 can therefore be regarded as fair representations of the intensity distribution curves obtained with the Fizeau interferometer using different beam dividers.

§ 7. CALCULATION OF FRINGE INTENSITY DISTRIBUTIONS

Using the general expression derived by Holden (1949) for the reflected fringe system from a parallel-plate interferometer, calculations were made of the theoretical fringe intensity distributions for the various beam-dividing surfaces used in this work. Figure 4 indicates the various optical processes involved, *a*, *b*, *c* and *d* representing the fractions by which unit amplitude of a

plane wave is modified in each case. The general expression takes account of the whole series of beams, progressively diminishing in intensity, which emerge from the interferometer after successive reflections at the opaque reflector and partial reflections and transmissions at the beam divider.

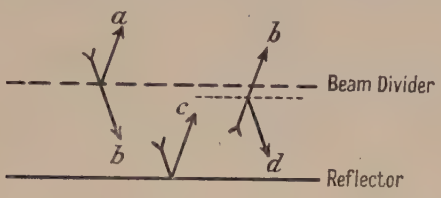


Figure 4. Optical processes in an interferometer.

For glass when uncoated, or when coated with a suitable film of titanium oxide, the calculations are much simplified because of the zero or negligible absorption of the dividing surface and the equality of its reflection factors a^2 and d^2. In these circumstances there is no apparent displacement of the reflecting planes for light incident on the 'back' and 'front' faces of the beam divider. Using numerical data for the various reflection and transmission factors at $\lambda = 0{\cdot}5\mu$ derived from Figure 1, and a reflection factor of $c^2 = 0{\cdot}4$ for lapped steel at normal incidence, the fringe intensity distributions shown in full line in Figure 5 were obtained. It will be seen that the calculated fringe shapes for a parallel-plate interferometer correspond very closely with those experimentally obtained from measurements of photographic density in Fizeau fringe images.

For the most suitable quality of aluminized beam divider used in gauge interferometry, a^2 differs from d^2 (see Figure 1) and the apparent positions of the reflecting planes for light incident on the 'back' and 'front' surfaces are not coincident. By suitable choice of a numerical value for the unknown phase constant thus involved (denoted by F in Holden's paper) and using values of the reflection and transmission factors at $\lambda = 0{\cdot}5\mu$ for aluminium taken from Figure 1, the asymmetrical fringe shape shown in dotted line in Figure 5 was obtained.

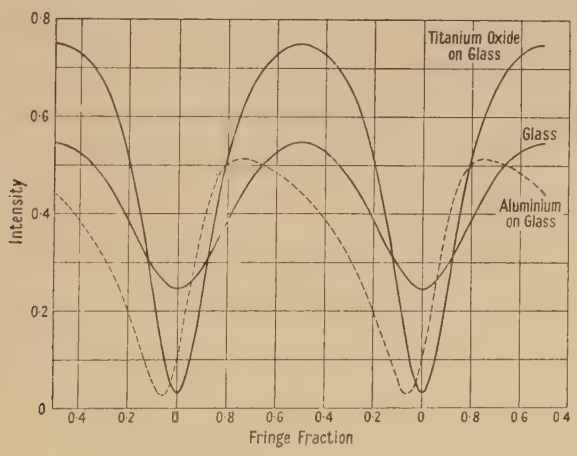


Figure 5. Calculated fringe intensity distributions for different beam dividers.

This agrees well with the shape derived experimentally from the density measurements plotted on Figure 3. The scale of fractional orders of interference in Figure 5 is reversed, i.e. decreasing from left to right, in order to display the asymmetry in the same sense as in Figure 3.

§8. CONCLUSIONS

Either Holden's formula or a simple direct calculation shows that, if all the multiple reflections are included, optimum contrast in the fringes is obtained when the reflection factor of the ideal non-absorbing beam divider is equal to that of the opaque reflector. For the special conditions existing in gauge interferometers a value of reflection factor of about 0·4 would therefore be preferable to the value of about 0·25 actually attained with the titanium oxide films. However, the same calculation shows that the improvement would be quite small, the contrast only changing slowly for reflection factors of the divider between about 0·25 and 0·55. There seemed little reason therefore to look for a suitable material giving a higher factor than the titanium oxide, or for adopting a multiple-film beam divider. Both from theoretical and practical considerations, therefore, a glass surface coated with a transparent titanium oxide film is demonstrated to be superior to beam-dividing surfaces hitherto used when associated with a surface of reflectivity 0·4 in Fizeau interferometers.

The results obtained also show that more than two beams may be allowed, with advantage, to become effective in fringe formation with the Fizeau instrument, even though the mirror separation is large.

ACKNOWLEDGMENTS

Acknowledgments are made to Mr. J. A. Hall for advice on the measurement and interpretation of photographic density, to Mr. H. G. W. Harding for measuring the spectral reflection and transmission curves of films and surfaces, and to Miss S. M. Jackson and Miss P. Teasdale-Buckell for assistance in the photographic and microphotometric work and for preparing the various records reproduced above.

The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

REFERENCES

Banning, M., 1947, *J. Opt. Soc. Amer.*, **37**, 688.
Barrell, H., 1948, *Research*, **1**, 533.
Cartwright, C. H., and Turner, A. F., 1939, *Phys. Rev.*, **55**, 1128.
Hall, J. A., 1949, *J. Sci. Instrum.*, **26**, 365.
Hammer, K., 1948, *Optik*, **3**, 495.
Hamy, M., 1906, *J. Phys. Radium*, **5**, 789.
Holden, J., 1949, *Proc. Phys. Soc.* B, **62**, 405.
Preston, J. S., 1950, *Proc. Roy. Soc.* A, **202**, 449.
Rouard, P., 1938, *Rev. Opt. Théor. Instrum.*, **17**, 1, 61, 89; 1949, *Ibid.*, **28**, 569.
Tolansky, S., 1948, *Multiple-Beam Interferometry of Surfaces and Films* (Oxford: Clarendon Press).

Fresnel Diffraction at a Transparent Lamina

By R. C. FAUST
British Rayon Research Association, Manchester

Communicated by H. G. Howell; *MS. received* 13*th June* 1950

ABSTRACT. Fresnel diffraction at the edge of a transparent lamina has been studied and an analogy drawn between the observed phenomena and interference fringes of the Fizeau and Edser–Butler types. The variation of the diffraction pattern with changing wavelength has been employed to determine accurately both the thickness and the refractive index of thin transparent sheets. The refractive index was found by an immersion method, the accuracy of which is here discussed. The Becke line which is often observed during such measurements is shown to be a strongly asymmetrical diffraction pattern occurring when the thickness of a lamina changes over a small but finite distance.

§1. INTRODUCTION

Some measurements of the refractive indices of transparent solids were recently made using a liquid immersion method, the immersed specimen being viewed under a microscope with a severely stopped-down condenser, thereby restricting the incident light to a narrow axial cone. This is the central illumination method (Saylor 1935) which has long been used for index determinations. It was then realized that the formation of the Becke line which is generally observed under these conditions cannot be adequately explained by a consideration of the refraction and internal reflection of the light at the specimen edge, but that a more rigorous approach was provided by a study of the Fresnel diffraction occurring at the edge. A systematic examination of such phenomena was therefore undertaken, transparent laminae being chosen as suitable test specimens.

When a parallel light beam falls upon a transparent sheet which has an abrupt and not a sloping or irregular edge, the observed Fresnel diffraction pattern is determined by the phase retardation δ suffered by the light upon traversing the sheet. At normal incidence $\delta=(2\pi/\lambda)(\mu-\mu_0)t$, where t is the sheet thickness, λ the vacuum wavelength of the light and μ and μ_0 are the appropriate refractive indices of the sheet and the surrounding medium. The intensity distribution in a plane parallel to that of the sheet can be calculated from the Kirchhoff diffraction theory; the resultant disturbance at any field point is that amplitude which would be found if the lamina were opaque plus the amplitude of the wave travelling through the lamina.

Such a calculation has been made by Kinder and Recknagel (1947). The incident parallel beam of light, which travels in the direction of the positive z-axis, falls upon the abrupt-edged sheet lying in the $x>0$ half of the xy plane. The objective is focused on the plane pp in front of the sheet, which as seen from Figure 1 (*a*) corresponds to positive z-values. The plane of observation is p′p′, the coordinates of an image point being so chosen as to agree with those of the corresponding point in the plane pp. In general the unit amplitude of the incident wave is weakened by the factor D upon travelling through the sheet, and

evaluation of the Kirchhoff integral leads to the following results for the disturbance u in the image plane p′p′.

$$\left.\begin{aligned} u &= 1 - A & x &< 0 \\ u &= De^{-i\delta} + A & x &> 0 \end{aligned}\right\} \qquad \ldots\ldots(1)$$

$$A = (1 - De^{-i\delta})\frac{e^{i\pi/4}}{\sqrt{2}}\int_{|\tau|}^{\infty} \exp\left(-\tfrac{1}{2}i\pi v^2\right) dv \qquad \ldots\ldots(2)$$

where $$\tau = -(2/\lambda z)^{1/2}x. \qquad \ldots\ldots(3)$$

The observed intensity I is obtained by multiplying u by its complex conjugate. If, as in Figure 1(*b*), the objective is focused upon a plane pp behind the sheet, z is

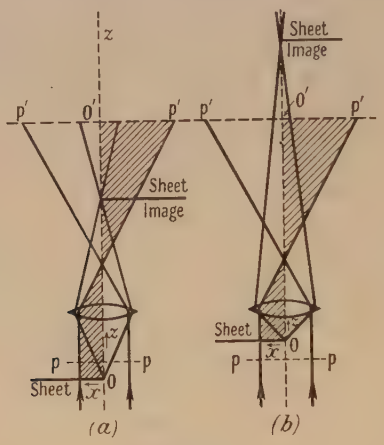


Figure 1.

negative and the diffraction pattern is virtual. Under these circumstances the observed intensity is given by the relationship

$$I(x, -|z|, \delta) = I(x, |z|, -\delta). \qquad \ldots\ldots(4)$$

The light intensity is a function of $x/(\lambda z)^{1/2}$, the surfaces of constant intensity being parabolic cylinders passing through O. The diffraction pattern at a given wavelength has therefore the same form on the set of planes z = constant, but the scale is proportional to $z^{1/2}$. The curves of Figure 2, reproduced from Kinder and

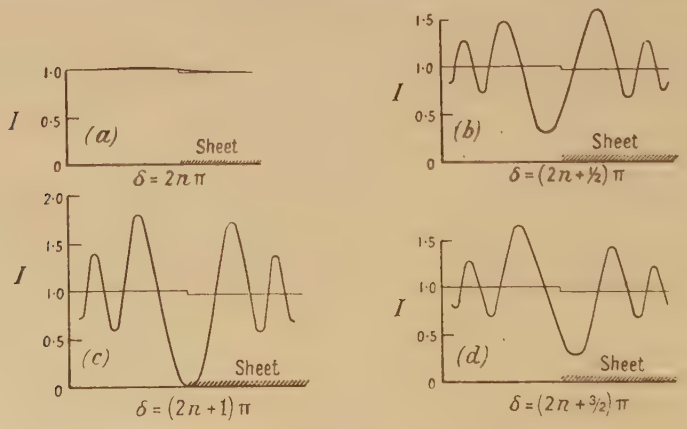


Figure 2.

Recknagel's paper, show the intensity distributions in the plane p′p′ for various phase retardations; D is taken as 0·975 and the plane pp lies in front of the sheet (Figure 1 (*a*)). When the phase retardation is an even multiple of π the pattern vanishes or has minimum visibility, whilst for odd multiples of π it is effectively symmetrical and of marked contrast. When $2n\pi<\delta<(2n+1)\pi$ the pronounced minimum lies outside the sheet, but when $(2n+1)\pi<\delta<(2n+2)\pi$ this minimum is within the sheet. Equation (4) shows that these asymmetrical forms are reversed if, as in Figure 1 (*b*), a plane pp behind the sheet is considered.

Several assumptions underlie this theoretical treatment: (i) The Kirchhoff theory is scalar, not vectorial, and ignores any polarization effects at large diffraction angles. In the series of experiments to be described no such effects were detected. (ii) The objective is aberration-free. (iii) Diffraction at the objective aperture is neglected. Using a 16 mm. objective (N.A. 0·3) high order diffraction bands have been clearly resolved although their separation was as small as the calculated resolution limit. Wegman (1950) has even observed bands separated by less than half the resolution limit. It would therefore appear that diffraction at the aperture of the imaging lens is generally of secondary importance.

§ 2. VARIATION OF SHEET THICKNESS

The condition that the diffraction pattern should possess minimum visibility is provided by the equation

$$n\lambda=(\mu-\mu_0)t. \qquad \ldots\ldots(5)$$

This is identical in form with that governing the transmission fringe positions arising from interference within a sheet, namely $n\lambda=2\mu t$. Thus the diffraction effects observed when the specimen thickness varies whilst μ, μ_0 and λ remain constant can be likened to Fizeau fringes: the disappearances of the diffraction pattern correspond to the fringe maxima and the symmetrical patterns of high contrast to the fringe minima. To demonstrate this a non-uniform piece of glass (about 0·04 mm. thick) was illuminated by a parallel beam of light (half-angle cone of 20′), the resulting diffraction pattern for 5000 A. being viewed with a 16 mm. objective raised by 2 mm. above the focal position. The glass film corresponds to the lower part of Figure 3 (Plate)*, the film thickness increasing towards the right. As expected from Figure 2 the pattern is periodic, the central minimum moving from inside to outside the film as one passes through a vanishing point in the direction of increasing thickness and phase retardation. On the negative more than thirty diffraction bands could be seen on either side of the edge.

When the objective is focused on the sheet $z=0$ the diffraction pattern should collapse and vanish. In practice slight departures from the ideal conditions theoretically postulated cause the specimen edge to appear as a narrow dark line on a bright background.

§ 3. VARIATION OF WAVELENGTH

Consider a transparent lamina illuminated with normally incident white light, the diffraction pattern of the edge being projected on to a spectroscope slit. The effect observed in the focal plane of the spectroscope is comparable with the well-known Edser–Butler interference fringes because the diffraction pattern generally disappears at a number of wavelengths throughout the spectrum, these disappearances corresponding to fringe maxima.

* For Plate see end of issue.

The wavelength variation of the diffraction pattern is dependent upon the dispersions of both the sheet and the surroundings, the effect of changing wavelength being given by the differential equation

$$(d\mu - d\mu_0)t = dn\lambda + nd\lambda. \qquad \ldots\ldots(6)$$

Applying the Cauchy dispersion formula one may express the relative change in refractive indices by

$$d\mu - d\mu_0 = -(A/\lambda^3)\omega d\lambda, \qquad \ldots\ldots(7)$$

where ω is the difference in the dispersions of the sheet and its surroundings, and A is a positive constant. One then obtains

$$d\lambda\{n + (A/\lambda^3)\omega t\} = -dn\lambda, \qquad \ldots\ldots(8)$$

from which it is seen that an increase in wavelength is always accompanied by a decrease in both the order and the phase retardation unless

$$-\omega > \frac{\lambda^3 n}{At} \quad \text{or} \quad -\omega > \frac{\lambda^2}{A}(\mu - \mu_0). \qquad \ldots\ldots(9)$$

If one assumes that the index of the sheet exceeds that of the ambient medium this inequality can only be satisfied when the dispersion of the sheet is sufficiently small relative to that of the medium. At 6500 A. the condition $-\omega > 0{\cdot}40(\mu - \mu_0)$ must be fulfilled, whilst at 4500 A. it is $-\omega > 0{\cdot}18(\mu - \mu_0)$.

Mica, which gives small abrupt cleavage steps, was chosen as a test material, the surrounding medium being air. The selected step showed disappearances of the pattern at 6730, 5430 and 4560 A., the step height being such that the product $(\mu - 1)t$ had the value $26{,}920 \pm 50$ A. at the wavelength 6730 A. Using a beam of plane-polarized light vibrating parallel to one of the principal axes of the mica an image of the diffraction pattern was cast on to a spectroscope slit by means of a 16 mm. objective. Figure 4 (Plate) shows the vanishing points at 5430 A. and 4560 A., the bottom half of the photograph corresponding to the higher level in the mica, that is, to the sheet. As the wavelength increases and δ decreases the pattern goes through those forms presented in Figure 2, i.e. as the wavelength increases beyond a vanishing point the pattern is initially asymmetrical, the chief minimum lying on the sheet side; perfect symmetry is then reached and as another vanishing point is approached the asymmetry reverses.

A sheet of mica was then immersed in a mixture of monobromonaphthalene and methyl salicylate; the index of the liquid was slightly less than that of mica but its dispersion was greater by 0·018. Under these conditions the inequality (9) was satisfied and the asymmetry was observed to be the opposite of that found with an air-surrounded film, the chief minimum lying on the film side when the wavelength was reduced below a vanishing point.

§4. APPLICATIONS OF FRESNEL DIFFRACTION

The above described behaviour of the diffraction pattern can be used to obtain an accurate evaluation of the product $(\mu - 1)t$, and consequently if μ is determined separately the specimen thickness t can be calculated. It will now be demonstrated that both of these quantities can be measured with the same experimental arrangements.

(*a*) *The Determination of Sheet Thickness*

The product $(\mu-1)t$ can be accurately measured for a sheet in air by noting the wavelengths at which the diffraction pattern disappears. This is most readily achieved by using a monochromator to illuminate the specimen with a parallel beam of light of continuously variable wavelength. From these wavelengths the order n can be found from the approximate equation $n\lambda_n=(n+p)\lambda_{n+p}$, where n and p are integers and $\lambda_n>\lambda_{n+p}$. Due to the dispersion of the material the calculated value of n will be slightly larger than its true integral value. The products $n\lambda_n \ldots . (n+p)\lambda_{n+p}$ are then plotted against $1/\lambda^2$, a straight line being drawn through the points.

Such a graph is presented in Figure 5 which applies to a mica step. As shown in Table 1 the positions of perfect symmetry (half-integral values for n) were noted as well as those of the vanishing points.

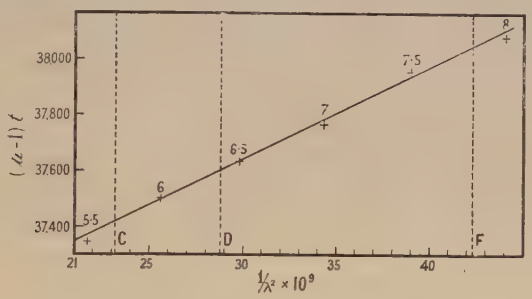


Figure 5.

Table 1

Wavelength (A.)	6790	6250	5790	5395	5060	4760
Order (n)	5·5	6·0	6·5	7·0	7·5	8·0
$n\lambda=(\mu-1)t$	37,345	37,500	37,635	37,765	37,950	38,080

All the points lie within 50A. of the line and the product $(\mu-1)t$ can therefore be read off at any wavelength to an even better accuracy. Under favourable circumstances the refractive index of a sheet of the thickness here considered (about 62,000 A.) can be determined to better than $\pm 0{\cdot}0003$, and as $\mu \simeq 1{\cdot}6$ this gives rise to an error of 1 in 2,000 in the value of $(\mu-1)$. Hence the final error in the thickness t should be 100 A. or less.

The slope of the line in Figure 5 also yields a value for the dispersion of mica. Taking the approximate refractive index as 1·6 the dispersion is found to be $0{\cdot}010 \pm 0{\cdot}002$.

Sheet thicknesses as great as 0·02 mm. can be measured to the same order of accuracy. The smallest sheet thicknesses which can be measured to this accuracy are those which produce a single vanishing point in the visible spectrum: i.e. those for which $(\mu-1)t$ exceeds 4000 A. If photography is used even thinner sheets can be studied. If only one vanishing point can be seen a decision regarding the order is made by noting the positions of perfect symmetry corresponding to half-integral values of n.

When no vanishing point can be found an estimate of the thickness follows from a qualitative examination of the diffraction pattern. If, as in Figure 2(*c*), the pattern is symmetrical and of good contrast then $(\mu-1)t$ is about $\frac{1}{2}\lambda$. If, however,

the pattern has the asymmetrical form of Figure 2(*b*), $(\mu-1)t$ has a value near $\frac{1}{4}\lambda$, whilst the opposite asymmetry of Figure 2(*d*) indicates an approximate value of $\frac{3}{4}\lambda$. Closer approximations should be possible by comparing at different wavelengths the patterns produced by the sheet under examination with those arising at a thicker sheet of known thickness.

The advantages of this method are that the specimen need not be mounted and that the attainable accuracy exceeds that possible with standard microscopy methods. The heights of coarse steps on crystal faces can also be measured readily, and furthermore one can ascertain which side of the crystal step is the more elevated by observing the asymmetry of the pattern as the wavelength passes through that corresponding to a vanishing point. Errors might be encountered, however, when dealing with high polymer sheets which are readily deformed, for example, cellulose acetate. If the specimen is cut from a larger sheet, the refractive index, the birefringence and the thickness at the edge will not necessarily be the same as those at the centre and some sectioning method must be used to provide an edge with a minimum of distortion. Such difficulties are also present when the standard methods are employed.

It is of interest to note that if the specimens are viewed in white light, those sheets which possess one or several vanishing points in the visible spectrum have brightly coloured diffraction patterns. If there is no vanishing point or if there are many, almost achromatic patterns are obtained.

(*b*) *The Determination of Refractive Index*

When a transparent solid is immersed in a liquid the diffraction pattern will vanish under the special condition that $\mu=\mu_0$. Such a disappearance of the image has long been used for index determinations.

It is generally stated that when using a stopped-down condenser a bright band will be observed to move across the specimen edge towards the medium of higher index if the objective is raised above the focal position. Lowering the objective causes this bright band, known as the Becke line, to travel towards the medium of lower index. In this manner one can discover whether to increase or decrease the index of the liquid in order to match it with that of the solid. When finally the matching is achieved no image of the specimen is visible. The writer has found, however, that such statements are only true when the specimen is relatively thick or has a sloping instead of an abrupt edge. A discussion of the experimental procedure and of the accuracy of the index measurements will therefore be deferred until the nature of the Becke line has been examined.

§ 5. ORIGIN OF THE BECKE LINE

The current explanation of the Becke line rests upon a consideration of the refraction and internal reflection of the light at the specimen edge. The only suggestion that diffraction might play a significant role appears to have been made by Spangenberg (1922) although nearly thirty years earlier Viola (1895) had remarked upon bands parallel to the specimen edges.

The origin of the Becke line can be discussed by examining the boundary wave concept advanced by Rubinowicz (1917), who has shown that the diffraction pattern due to a screen can be calculated by considering that a wave spreads out from the screen edge, the resultant amplitude at any point being the amplitude

of the boundary wave plus that which would obtain if the light propagation were strictly linear. Figure 1 shows the boundary wave centred at the edge of the transparent lamina. If the wave is symmetrical about the z-axis this treatment leads to essentially the same results as obtained by Kinder and Recknagel from the Kirchhoff theory but, as will now be shown, any asymmetry tends to produce a Becke line.

Assume that the refractive index of the sheet is greater than that of the surroundings and that the amplitude of the boundary wave is greater for positive values of x than for negative values. If the objective be raised above the focal position (Figure 1 (*a*)) the interference occurring to the right of the point O′ on the image plane takes place between the rectilinearly propagated light and a strong boundary wave; to the left of O′, however, the interference effects will be relatively slight. The diffraction pattern is therefore more marked on the sheet side of the edge, and with strong asymmetry the pattern assumes the form of the Becke line. Similarly when the objective is lowered interference is more marked to the left of O′, i.e. outside the sheet.

Martin (1943) studied experimentally the boundary wave arising at the edge of a transparent lamina of refractive index higher than that of the surroundings. If the lamina edge was not abrupt but changed its thickness over a small distance, the boundary wave amplitude was found to be greater on the sheet side of the edge (i.e. for positive values of x), which is exactly the assumption made in the preceding paragraph. He also predicted this asymmetry from an examination of the Cornu spiral.

Thus, if a thin specimen with an abrupt edge is illuminated with a parallel beam of light at normal incidence, the theory of previous sections can be applied. The diffraction pattern is then clearly visible on both sides of the specimen edge and one cannot in general decide whether the solid or the liquid has the higher index. On the other hand a specimen with a slightly tapering edge produces a diffraction pattern which is not equally marked on the two sides of the boundary; if the objective is raised above the focal position the maxima and minima are stronger on that side of the edge corresponding to the optically denser medium. Such a diffraction pattern, which is represented by the vertical diffraction bands in Figure 6 (Plate), was observed by immersing a glass specimen, 0·05 mm. thick, in a liquid of slightly higher index and illuminating it with a parallel beam of monochromatic radiation. With an even less ideal specimen the diffraction pattern is almost completely lost on the low index side, whilst on the high index side one perceives a very bright maximum sometimes bordered by weak bands. This is the Becke line which is to be interpreted as the first bright band of an asymmetrical diffraction pattern. It is seen along the horizontal edge of Figure 6.

It is to be noted that the boundary wave arising at an abrupt edge will be asymmetrical if the incident beam travels at an angle to the z-axis, thus explaining the enhanced asymmetry produced by the use of a narrow-aperture condenser. Under such conditions the transition region over which the optical retardation alters will increase with specimen thickness, thereby leading to readier formation of the Becke line. As, however, the light from each element of the source is incoherent the intensities are additive and, if the range of incident angles passed by the condenser is large, the asymmetry will be partly destroyed. If the numerical aperture of the objective is too small to resolve the region over which the sheet thickness changes, a single boundary wave equivalent to that arising

at an abrupt edge is involved. If, however, the transition region is resolved by the objective each point of the tapering sheet produces its own boundary wave, the phase and amplitude of which corresponds to that of a sheet of smaller thickness than is actually employed. The sensitivity of the central illumination method should therefore suffer as the numerical aperture of the objective is increased. This has been confirmed, as is described in the following section.

§ 6. THE CENTRAL ILLUMINATION METHOD

Of the various means available for the determination of refractive indices by the immersion method, that depending upon the different dispersions of the solid and the liquid was employed (Posnjak and Merwin 1922). Two non-volatile miscible liquids are stirred together on a cavity slide to form a stable mixture whose index is close to that of the solid. Using a monochromator the defocused image of the immersed specimen is examined until a wavelength is found at which the image vanishes, i.e. for which $\mu = \mu_0$. At a given wavelength one can decide whether the solid or the liquid possesses the greater index by noting the position of the Becke line which tends to appear with all specimens except those thin ones with truly abrupt edges. From the determined dispersion curve of the liquid the index of the solid can be found for that wavelength corresponding to the vanishing point. The composition of the liquid mixture is now changed slightly by adding a drop or two of one of the liquids with a hypodermic syringe. A new vanishing point is thereby obtained and repetition of the process permits one to draw a dispersion curve for the solid. Dispersion curves are most conveniently drawn by plotting μ against $1/\lambda^2$ to produce a straight line.

In practice the image does not vanish at a single wavelength but disappears over a wavelength range of half-width $d\lambda$. The results shown in Table 2 were obtained with glass films immersed in mixtures of methyl and ethyl salicylates, the measurements, which were made with a 16 mm. objective without a condenser, being confined to steep edges. The difference between the dispersions of the solid and liquid was $\omega = -0{\cdot}01$, and the least detectable relative change $(d\mu - d\mu_0)$ in the indices was calculated from equation (7).

Table 2

	$d\lambda$ (A.)		$(d\mu - d\mu_0)$		$dn\lambda$ (A.)	
t (mm.) $\times 10^{-3}$	6000 A.	5000 A.	6000 A. $\times 10^{-5}$	5000 A. $\times 10^{-5}$	6000 A.	5000 A.
5	190	100	91	83	45	41
6·5	180	70	87	58	57	38
12	90	37	43	31	52	37
41	21	12	10	10	41	41
118	13	7	6	6	53	49
164	9	4	4	3	70	54

Noting that the order is zero when $\mu = \mu_0$, one obtains from equation (6)

$$(d\mu - d\mu_0)t = dn\lambda.$$

If it is assumed that a change of order dn produces an image which is just perceptible, then for a given wavelength the product $(d\mu - d\mu_0)t$ should be constant and the smallest detectable change in the relative indices should diminish with an increase

of specimen thickness. This is borne out by the experimental figures for the two wavelengths 6000 A. and 5000 A. Furthermore the change of order required to give a perceptible image is almost the same for both wavelengths, namely $dn = 0{\cdot}01$.

These results suggest a measure for the resolution in depth of a microscope, a step being visible if $(\mu - \mu_0)t = 0{\cdot}01\lambda$. If $\mu = 1{\cdot}5$ and $\mu_0 = 1$ a step as small as 100 A. high can be detected using green light. If the step is viewed in reflection the height necessary for detection is given by $2\mu_0 t = 0{\cdot}01\lambda$, i.e. 25 A. Unlike the resolution limit, which is smaller the greater the numerical aperture of the objective, a change of objective exerts little influence over the resolution in depth, nor does the use of a stopped-down condenser materially affect the value.

So far the edge has been regarded as abrupt. If it is sloping the wavelength range over which the specimen vanishes is increased and shows a marked dependence upon the numerical aperture (Table 3). This confirms the conclusion drawn by Saylor that the central illumination method is less sensitive with high-power objectives and with specimens possessing sloping edges.

Table 3

	N.A.	$d\lambda$(6000 A.)	
40 mm. × 10 ocular	0·15	17±5	16±4
16 mm. × 10 ocular	0·30	54±7	30±5
8 mm. × 5 ocular	0·65	57±10	39±6
Specimen thickness (mm. × 10^{-3})		118	164

Although the image disappearance with thin specimens covers a wide wavelength band, the refractive index can be measured very accurately by employing the mean of the two limiting wavelengths at which the image is just visible. For example, the individual index measurements made upon a glass film, 0·005 mm. thick, deviated from a straight line dispersion curve by no more than ±0·0002. Furthermore if the incident beam is a parallel one, strictly plane-polarized light can be obtained and consequently any measurements made on highly birefringent substances will be more accurate than those found when a condenser is introduced or when an oblique incidence method is employed.

ACKNOWLEDGMENTS

The author wishes to express his thanks to Dr. H. G. Howell for helpful discussions during the course of this study and to Mr. B. J. Corr for assistance with the photography. This work forms part of the programme of fundamental research undertaken by the British Rayon Research Association.

REFERENCES

Kinder, E., and Recknagel, A., 1947, *Optik*, **2**, 346.
Martin, L. C., 1943, *Proc. Phys. Soc.*, **55**, 104.
Posnjak, E., and Merwin, H. E., 1922, *J. Amer. Chem. Soc.*, **44**, 1970.
Rubinowicz, A., 1917, *Ann. Phys., Lpz.*, **53**, 257.
Saylor, C. P., 1935, *J. Res. Nat. Bur. Stand., Wash.*, **15**, 277.
Spangenberg, K., 1922, *Fortschr. Min.*, **7–8**, 1.
Viola, C., 1895, *Tschermak's Min. Petr. M.itt*, **14**, 554.
Wegman, L., 1950, *Helv. phys. Acta*, **23**, 437.

The Refractive Index of Electron Optics and its Connection with the Routhian Function

By W. GLASER
Institute for Applied Physics, Technische Hochschule, Vienna

Communicated by V. E. Cosslett; MS. received 12*th June* 1950 *and, in final form* 16*th October* 1950

ABSTRACT. The derivation of the electron-optical refractive index from Hamilton's Principle is discussed. It is shown that the criticisms made by Ehrenberg and Siday of our method are invalid, and that our derivation made in 1933 is correct. To prove that our method is generally applicable, the isotropic refractive index of an axially symmetrical field is obtained.

In the theory of electron optics, it has proved useful in many ways to determine the geometrical form of the electron paths in an electromagnetic field by the use of Fermat's principle (see, for instance, Cosslett 1946, Busch and Brüche 1937). Apart from its fundamental importance that it connects the method of electron optics with those of ordinary optics, the specification of electron paths by a variational equation (Schwarzschild 1903) is valuable in that it does not introduce a particular coordinate system. Furthermore, this formulation lends itself to a straightforward calculation of aberrations and, finally, facilitates the direct approach to the wave-mechanical treatment of electron optics. We first introduced this method in 1933 in electron optics and gave the refractive index corresponding to an electromagnetic field (Glaser 1933). In a paper entitled 'The Refractive Index in Electron Optics and the Principles of Dynamics', by Ehrenberg and Siday (1949) which was recently published in this journal, the authors remark that our method, although it leads to correct results, is not correct in its derivation. We propose, in the following, to show that our derivation of the electron-optical refractive index is correct and that the criticism of the two authors mentioned is based on a misconception.

As is well known (Nordheim and Fues 1927), the transition from Hamilton's Principle $\delta\int L\,dt=0$, which is time-dependent, to Maupertuis' Principle of Least Action or to Fermat's Principle $\delta\int\mu\,ds=0$, cannot be performed simply by writing $\mu=L/v$. In this, everyone will agree with the two authors. However, they have overlooked the fact that we did not proceed in this way in our work; in deriving μ, we added the term E/v to $\mathscr{L}$, where $\mathscr{L}=L/v$, as is necessary in reducing the variational principle $\delta\int L\,dt=0$ by means of the energy integral. E stands for the total energy which Ehrenberg and Siday denoted by $-h$. We wish now to consider the question very briefly, and show how our method can also be successfully applied to obtain a simplified refractive index when an integral of the equations of motion is known.

First of all, it is obvious from the formulation of the two principles that our derivation—when one considers the addition of E—must always lead to the same result as that given in the paper cited; indeed, the two methods are identical. The energy integral, if L is not explicitly dependent on time, i.e. $\partial L/\partial t=0$ has the form

$$\Sigma\dot{q}_k\frac{\partial L}{\partial\dot{q}_k}-L=E, \qquad \ldots\ldots(1)$$

Hence the refractive index

$$\mu = \frac{L+E}{v} = \frac{1}{v}\Sigma \dot{q}_k \frac{\partial L}{\partial \dot{q}_k}, \qquad \ldots\ldots(2)$$

which is our expression. Thus the validity of our results is in no way 'fortuitous', as has been said. The addition of E to L is, however, essential as we shall soon see.

The shortest way to prove and, at the same time, to generalize our method is to use the well known Routhian equation of motion and apply it to cyclic systems (Routh 1877, Whittaker 1904, Frank–Mises 1935). Wallauschek (1937) in his thesis gave an application of this equation of motion treating a mass spectrograph with double direction focusing (cf. Svartholm and Siegbahn (1946)). The Routhian equations of motion are something between the Hamilton and the Lagrange equations of motion and lead directly to the reduction of the variational problem to one in a smaller number of variables if some of the coordinates are cyclic or negligible, i.e. if they do not occur explicitly in the Lagrangian function. Let us assume that

$$\delta \int \mathscr{L}(\dot{q}_k, q_k')\, du = 0 \qquad \text{wherein} \qquad q_k' = \frac{dq_k}{du}, \qquad \ldots\ldots(3)$$

the variable u and the magnitudes of the q_k at the end-points not being subject to variation. We use u rather than t as the parameter for reasons which will become apparent later. The Euler–Lagrange equations derivable from (3) are

$$\frac{d}{du}\frac{\partial \mathscr{L}}{\partial \dot{q}_k} - \frac{\partial \mathscr{L}}{\partial q_k} = 0. \qquad \ldots\ldots(4)$$

We can obtain a different form for the equations of motion if we separate the n coordinates $q_1, q_2, q_3 \ldots q_n$ into two groups

$$q_1, q_2, \ldots q_\mu, \; q_{\mu+1} = \rho_1, \qquad q_{\mu+2} = \rho_2 \ldots q_n = \rho_r,$$

where $\mu + r = n$, and eliminate the velocities ρ_k' which belong to the second group by means of the equations

$$\pi_k = \frac{\partial \mathscr{L}}{\partial \rho_k'} \qquad k = 1, 2 \ldots r. \qquad \ldots\ldots(5)$$

All functions of the n coordinates and velocities $q_1, q_2 \ldots q_n, q_1', q_2', \ldots q_n'$ can be expressed as functions of the quantities $q_1, q_2 \ldots q_\mu, q_1' \ldots q_\mu', \rho_1, \rho_2 \ldots \rho_n, \pi_1, \pi_2, \ldots \pi_r$. One introduces to this end a function which is something between the Hamiltonian and Lagrangian functions, namely the Routhian function R defined by

$$R(q_1, \ldots q_\mu, q_1' \ldots q_\mu', \rho_1 \ldots \rho_r, \pi_1 \ldots \pi_r) = -\mathscr{L} + \sum_{k=1}^{r} \pi_k \rho_k'. \qquad \ldots\ldots(6)$$

It can then easily be shown that the equations of motions (4) fall into the following two groups

$$\frac{d}{du}\frac{\partial R}{\partial q_k'} - \frac{\partial R}{\partial q_k} = 0, \qquad k = 1, 2 \ldots \mu; \qquad \ldots\ldots(7)$$

and

$$\frac{\partial \rho_j}{du} = \frac{\partial R}{\partial \pi_j}, \qquad \frac{d\pi_j}{du} = -\frac{\partial R}{\partial \rho_j}, \qquad j = 1, 2 \ldots r. \qquad \ldots\ldots(8)$$

We shall now assume that the coordinates of the second group, which we called $\rho_1, \rho_2 \ldots \rho_r$ are cyclic, i.e. they do not appear explicitly in the function $\mathscr{L}$. Then R is also independent of $\rho_1, \rho_2 \ldots \rho_r$ and it follows from (8) that the corresponding

'momenta' $\pi_1, \pi_2 \ldots \pi_j$ are all constant. They therefore only appear as constant parameters in the function R and in the equations of motion (7). These, however, are the Euler–Lagrange equations derivable from the Lagrangian function (6), i.e. from the variational equation

$$\delta \int (-\mathscr{L} + \Sigma \pi_k \rho_k') \, du = 0. \qquad \ldots\ldots (9)$$

Since the integrand does not involve the second group of variables the equations of motion will hold, even if these variables are subject to variation at the limits of integration. We shall give two applications in the field of electron optics: (i) Derivation of the refractive index of an arbitrary electromagnetic field; this will prove to be identical with our derivation in 1933. (ii) Establishment of the isotropic refractive index of axially symmetrical fields; this shows that our method is capable of generalization.*

(i) *Refractive Index of an Arbitrary Electromagnetic Field*

Let the electron paths be represented by

$$\delta \int L(q_k, \dot{q}_k, t) \, dt = 0. \qquad \ldots\ldots (10)$$

The coordinates q_k and the time t are not to be varied at the limits. We introduce a new parameter u, by setting

$$t = t(u), \quad q_k = q_k(u); \qquad k = 1, 2 \ldots n. \qquad \ldots\ldots (11)$$

By this means we separate the motion into the geometrical path $q_k = q_k(u)$, $k = 1, 2 \ldots n$ and the time $t = t(u)$. We write

$$\delta \int L\left(q_k, \frac{\dot{q}_k}{t'}, t\right) t' \, du = 0 \quad \text{with} \quad q_k' = dq_k/du, \quad t' = dt/du. \qquad \ldots\ldots (12)$$

If we set

$$\mathscr{L}(q_k, q_k', t) = L(q_k, q_k'/t, t)t' \qquad \ldots\ldots (13)$$

we obtain from

$$\delta \int \mathscr{L} \, du = 0, \qquad \ldots\ldots (14)$$

the Lagrange equations of motion

$$\frac{d}{du}\frac{\partial \mathscr{L}}{\partial \dot{q}_k} - \frac{\partial \mathscr{L}}{\partial q_k} = 0, \quad \text{or} \quad \frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{q}_k} - \frac{\partial \mathscr{L}}{\partial q_k} = 0, \qquad \ldots\ldots (15)$$

and

$$\frac{d}{du}\frac{\partial \mathscr{L}}{\partial t'} - \frac{\partial \mathscr{L}}{\partial t} = 0 \qquad \ldots\ldots (16)$$

if we do not vary q_k and t at the limits.

Equation (16), which results from variation of the time $t = t(u)$, may be transformed by means of

$$\frac{\partial \mathscr{L}}{\partial t'} = L - \Sigma q_k \frac{\partial \mathscr{L}}{\partial \dot{q}_k} \qquad \ldots\ldots (17)$$

into the well-known identity

$$\frac{d}{dt}\left(L - \Sigma \dot{q}_k \frac{\partial L}{\partial \dot{q}_k}\right) = \frac{\partial L}{\partial t}, \qquad \ldots\ldots (18)$$

which can also be deduced from the equations of motion (15). However, if the time is not contained in $\mathscr{L}$ or in L, then from (16)

$$\partial \mathscr{L}/\partial t' = \pi = -E \qquad \ldots\ldots (19)$$

* This has been challenged by Ehrenberg and Siday with reference to a paper by Opatowsky.

in which the constant π, which denotes the negative energy, is given by $-E$. The required new variational principle, in which time is no longer contained, follows from (9)

$$-\delta \int R\,du = -\delta \int (\mathscr{L} - \pi t')\,du = -\delta \int \frac{L+E}{v}\,ds = 0, \quad \ldots\ldots(20)$$

where we have specified $u = s$, that is $t = 1/v$.

This is, however, precisely the form of the refractive index that was obtained in our 1933 paper and the method by which it was derived. For from

$$L = \tfrac{1}{2}mv^2 + e\phi - e(\mathbf{Av}) \quad \ldots\ldots(21)$$

and

$$\tfrac{1}{2}mv^2 - e\phi = E \quad \ldots\ldots(22)$$

it follows that

$$\mu = \frac{L+E}{v} = \frac{1}{v}[mv^2 - e(\mathbf{Av})] = mv - e(\mathbf{As}), \quad \text{if} \quad \mathbf{s} = \frac{1}{v}\mathbf{v} \quad \ldots\ldots(23)$$

for which we may also write, with the generalized impulse $\mathbf{g} = m\mathbf{v} - e\mathbf{A}$

$$\mu = \mathbf{gs} = p_s. \quad \ldots\ldots(24)$$

The velocity is in this case to be expressed as a function of position according to (22).

In the special case where the potential is so normalized that $E = 0$, we can immediately write for the refractive index

$$\mu = L/v \quad \ldots\ldots(25)$$

eliminating the velocity by means of the energy relations

$$\tfrac{1}{2}mv^2 + U = 0; \qquad mv = (-2mU)^{1/2}. \quad \ldots\ldots(26)$$

For instance, if the zero of electric potential is chosen so that it vanishes when the electron velocity is zero, as is often convenient, then $\frac{1}{2}mv^2 - e\phi = 0$ is valid, i.e. $E = 0$. In this case the index of refraction is obtained at once from $\mu = L/v = (1/v)(\frac{1}{2}mv^2 + e\phi - e\mathbf{Av})$ by replacing mv by $(2me\phi)^{1/2}$ and $\mathbf{v}/v$ by $\mathbf{s}$.

(ii) *The Isotropic Refractive Index of Axially Symmetrical Fields*

Assuming that the electromagnetic field is rotationally symmetrical, we have for the element of arc length in cylindrical coordinates

$$ds = (dr^2 + r^2\,d\theta^2 + dz^2)^{1/2}. \quad \ldots\ldots(27)$$

For $\mu\,ds$, since $A_r = 0$, $A_z = 0$, $A_\theta = A$ in conditions of rotational symmetry, we have

$$\mu\,ds = mv\,ds - eAr\,d\theta; \qquad \tfrac{1}{2}mv^2 - e\phi = E(=0). \quad \ldots\ldots(28)$$

From (27) and (28) it follows that

$$\mu\,ds = [\{2me\phi(r'^2 + r^2\theta'^2 + z'^2)\}^{1/2} - eAr\theta']\,du \quad \ldots\ldots(29)$$

with $r' = dr/du$, etc. In the Lagrange function

$$\mathscr{L} = \{2me\phi(r'^2 + r^2\theta'^2 + z'^2)\}^{1/2} - eAr\theta', \quad \ldots\ldots(30)$$

θ is therefore a cyclic variable; this expresses the rotational invariance of the problem (axial symmetry). Since $\partial\mathscr{L}/\partial\theta = 0$ it follows that

$$\frac{\partial\mathscr{L}}{\partial\theta'} = \pi = \text{const.} = (2me\phi)^{1/2}\frac{r^2\theta'}{(r'^2 + r^2\theta'^2 + z'^2)^{1/2}} - eAr \quad \ldots\ldots(31)$$

and the Routhian function becomes

$$R = -\mathscr{L} + \pi\theta', \qquad \ldots\ldots(32)$$

where θ' is to be taken from equation (31). If we then similarly substitute for θ' in

$$-R = \{2me\phi(r'^2 + r^2\theta'^2 + z'^2)\}^{1/2} - (\pi + Ar)\theta', \qquad \ldots\ldots(33)$$

we obtain

$$-R = [2me\phi - \{(\pi/r) + eA\}^2]^{1/2}(r'^2 + z'^2)^{1/2}, \qquad \ldots\ldots(34)$$

and the electron path may be represented by

$$-\delta \int R\,du = \delta \int [2me\phi - \{(\pi/r) + eA\}^2]^{1/2} d\sigma \text{ with } d\sigma = (dr^2 + dz^2)^{1/2}. \qquad \ldots\ldots(35)$$

The 'meridional' motion thus takes place in the r, z plane (which rotates according to (31)), corresponding to the Fermat principle, as in an optical medium of refractive index

$$\mu_{is} = [2me\phi - \{(\pi/r) + eA\}^2]^{1/2}. \qquad \ldots\ldots(36)$$

That the meridional motion may be expressed as a potential motion in a potential field

$$Q = e\phi - \frac{1}{2m}\left(\frac{\pi}{r} + eA\right)^2 \qquad \ldots\ldots(37)$$

was shown by Störmer (1907) and later independently by Busch (1927), by direct evaluation of electron paths in axially symmetrical fields. The above direct derivation shows how this fact is necessarily connected with the axial symmetry of the general index of refraction. It also shows that the Routhian method, as used by us to reduce the number of variables in the variational problem, can be successfully applied to other cases, in opposition to the views of Ehrenberg and Siday.

REFERENCES

Busch, H., 1927, *Arch. Elektrotech.*, **18,** 583.

Busch, H., and Brüche, E., 1937, *Beiträge zur Elektronenoptik* (Leipzig : Barth), p. 24.

Cosslett, V. E., 1946, *Introduction to Electron Optics* (Oxford : University Press), p. 249.

Ehrenberg, W., and Siday, R. E., 1949, *Proc. Phys. Soc.* B, **62,** 8.

Frank, Ph., v. Mises, R., 1935, *Die Differential- und Integralgleichungen der Mechanik und Physik*, vol. 1 (Brunswick : Vieweg), p. 59.

Glaser, W., 1933, *Z. Phys.*, **81,** 451.

Nordheim, L., and Fues, E., 1927, *Handbuch der Physik von Geiger und Scheel*, vol. 5 (Berlin : Julius Springer), chap. 3.

Routh, E. J., 1877, *A Treatise on the Stability of a given State of Motion* (London : Macmillan).

Schwarzschild, K., 1903, *Nachr. Ges. Wiss. Göttingen*, **3,** 126.

Störmer, C., 1907, *Arch. Sci. phys. nat., Geneva*, 1907.

Svartholm, N., and Siegbahn, K., 1946, *Ark. Astr. Phys.*, **33A,** No. 21.

Wallauschek, R., 1937, *Thesis* (Prague).

Whittaker, E. T., 1904, *A Treatise on the Analytical Dynamics* (Cambridge : University Press), pp. 54–55.

Stress Birefringence in Polyethylene

By S. M. CRAWFORD and H. KOLSKY
Imperial Chemical Industries Limited, Butterwick Research Laboratories, Welwyn, Herts.

MS. received 11*th July* 1950, *and in final form* 25*th August* 1950

ABSTRACT. The birefringence and strain produced in polyethylene by different stresses have been measured at temperatures between 20° and 90° C. It has been found that the relation between strain and birefringence is linear up to strains of 0·25 and for greater strains there is still no hysteresis. This indicates that no true flow occurs in this temperature range. The stress–birefringence relation, on the other hand, rapidly departs from linearity and shows large hysteresis, so that the nature of the birefringence effects is in direct contrast with that observed in rubber. An expression has been derived for the birefringence in terms of the orientation of rod-like crystallites embedded in an elastic matrix and this was found to be in qualitative agreement with the experimental results.

§ 1. INTRODUCTION

Earlier work (Kolsky and Shearman 1943) has shown that the birefringence produced in polyethylene when it is stretched is in agreement with the molecular re-arrangements deduced from chemical and x-ray evidence. This optical technique is especially valuable in studying low degrees of orientation where x-ray diffraction is not sufficiently sensitive to show any effect. Bunn (1941) has found from x-ray evidence that polyethylene is partly amorphous and partly crystalline, the crystallites being small compared with the molecular chain length so that each molecule goes through several crystal groupings.

The earlier results have shown that for very small stresses, both the strain and the birefringence are proportional to the stress. At higher stresses, large hysteresis effects were observed and when finally the specimen was *cold drawn*, removal of the load resulted in practically no decrease in either the strain or the birefringence.

In the present investigation the birefringence and strain produced on stressing polyethylene at different temperatures have been measured. The purpose was to see whether any true flow takes place on stretching or if a reversible orientation of the long chain molecules is all that occurs. An attempt has also been made to interpret the strain–birefringence curves quantitatively in terms of the orientation of the crystallites on stretching.

§ 2. EXPERIMENTAL

The optical arrangement is shown in Figure 1. This is similar to that used in the earlier work, except that the strain was observed at the same time as the birefringence by measuring the distance between the images of two lines on the specimen with a cathetometer. The setting of the mirror in this arrangement was not critical since all that was required was a ratio of distances.

The stresses were applied by a pulley and weights and provision was made for moving the lower end of the specimen to ensure that the light always passed through the portion between the two marked lines. The Babinet compensator was calibrated for mercury green light ($\lambda = 5461$ A.) and used as in the earlier work, the path difference being measured to about one-hundredth of a wavelength. The

birefringence Δn is given as the difference between the refractive indices for light of this wavelength polarized along and perpendicular to the direction of stretching. For the thickness of specimens used, this corresponds to an accuracy within 1×10^{-4} in the value of Δn.

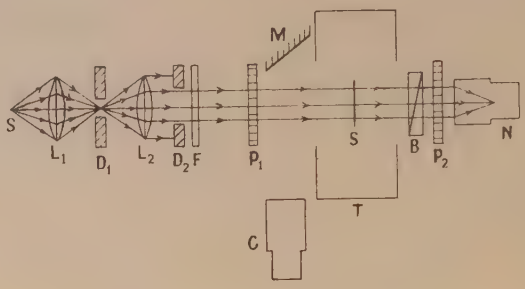


Figure 1. Experimental arrangement.
S=source, L_1L_2=lenses; D_1D_2=diaphragms; F=filter; P_1P_2=polaroid screens; M=plane mirror; C=cathetometer; T=constant temperature enclosure; S=specimen; B=Babinet compensator; N=microscope.

Rectangular specimens 5 cm. × 2 cm. were cut from sheets of polyethylene 0·012 cm. thick, which had been carefully annealed to remove any initial strain. These were taken through hysteresis cycles at various temperatures between 20° and 90° c. and readings of strain and birefringence were recorded at each value of the stress.

§3. RESULTS

The most detailed experiments were carried out for strains up to 0·1 and Figures 2, 3, 4 show the results obtained at different temperatures. Figures 2 and 3

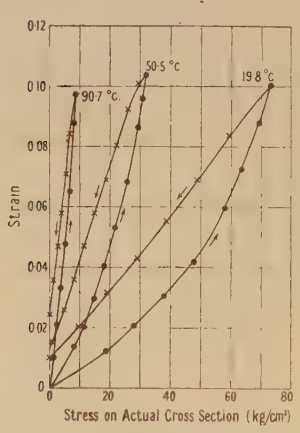


Figure 2.

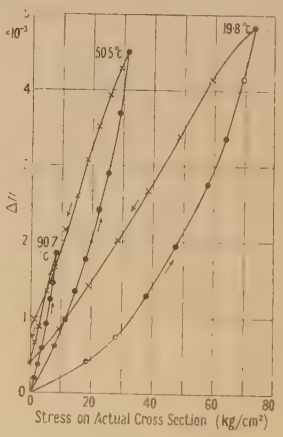


Figure 3.

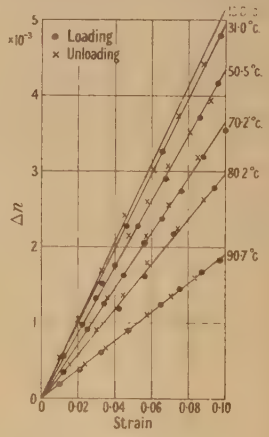


Figure 4. For clarity, no experimental points are shown at 19·8°c.

show graphs of the strain and birefringence with stress. It can be seen in both cases that the curves depart from linearity at very low stresses and the rate of rise increases with temperature. On reducing the stress, large hysteresis effects are observed. In Figure 4, the birefringence is plotted against strain, and here the relation is linear in each case, the rate of rise decreasing with rise in temperature. Figure 5 shows the gradient of the (Δn, strain) curve plotted against temperature.

This indicates that it should reach zero at 110° c., which is in agreement with results for the *softening point* of the material (Richards 1945).

For temperatures up to 50° c., results could be obtained for greater strains, and Figure 6 shows the (Δn, strain) curves obtained under these conditions. It may be seen that although the curves depart from linearity above a strain of about 0·25, there still appears to be no hysteresis on unloading.

These results show that the birefringence is a univalued function of strain—at least up to temperatures of 90° c.—and it is only possible to remove any orientation present by allowing corresponding changes in the shape of the specimen. In order to see whether this was still true for longer periods of time, specimens were clamped at constant strain and the birefringence measured over a period of time. Figure 7 gives the results for four specimens at different temperatures.

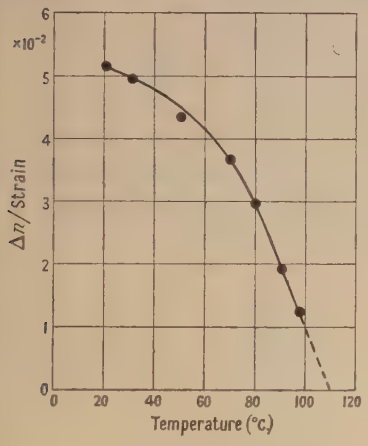


Figure 5.

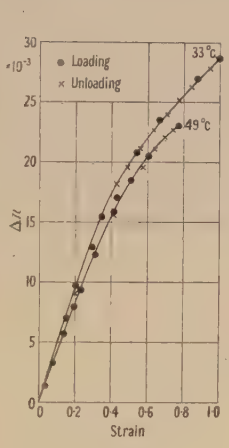


Figure 6.

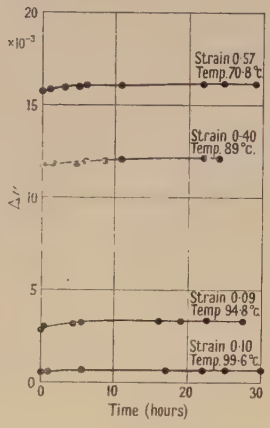


Figure 7. Time dependence of birefringence at constant strain.

These show that the birefringence *increases* slightly with time, so that attempts to anneal specimens under constant strain at these temperatures would appear, in fact, to increase the orientation. This gradual increase in the orientation may be associated with the growth of crystals along the direction of stretching.

If the specimen is allowed to contract freely, the strain and birefringence decrease when the temperature is raised. Figures 8 and 9 show the effect of carrying this out with specimens of different initial strain. For low strains, the greater part both of the strain and of the birefringence has disappeared at a temperature of 50° c., and any further increase in temperature has a smaller effect. In the case of the higher initial strains both the strain and birefringence appear to fall steadily up to 100° c. This indicates that the thermal energy necessary to disorient the molecular groupings is greater once they have become aligned along the direction of stretching. This type of behaviour might be expected from the fact that when polyethylene is cold-drawn, it shows practically no retraction on releasing the stress, suggesting that the oriented crystallites form a stable structure.

§ 4. DISCUSSION

The fact that at any one temperature in the range considered, the birefringence of polyethylene is a univalued function of strain and not of stress indicates that

it is produced largely by the geometrical orientation of the crystallites rather than by distortion due to the stress. The orientation arising from purely geometrical considerations is discussed in the Appendix. If it is assumed that the crystallites are uni-axial, rod-shaped and embedded in an elastic matrix, the birefringence is given by

$$\Delta n = \frac{k}{\pi}\int_{\theta=0}^{\theta=\pi/2}\int_{\varphi=0}^{\varphi=\pi}\frac{\tan\theta\sec^2\theta}{(k^2+\tan^2\theta)^{3/2}}\left\{\frac{n_2}{(1+m\sin^2\theta\sin^2\phi)^{1/2}}-n_1\right\}$$
$$\times\left\{\frac{2\cos^2\theta}{1-\sin^2\theta\sin^2\phi}-1\right\}d\theta\,d\phi,$$

where n_1, n_2 are the ordinary and extraordinary refractive indices respectively, $m=(n_2^2-n_1^2)/n_1^2$ and k is the ratio of the width to the length of a square of material after stretching. When n_2-n_1 is small compared with unity, the value of this integral approximates to

$$\frac{n_2-n_1}{2}\left\{\frac{3}{(1-k^2)}-\frac{3k\cos^{-1}k}{(1-k^2)^{3/2}}-1\right\}.$$

When $k\to 0$ the expression $\to(n_2-n_1)$. This value corresponds to perfectly oriented material and the earlier work (Kolsky and Shearman 1943) has shown that its limiting value is 0·044. This value agrees with that given by Bunn (1949) and with that found for single crystals of long chain paraffins.

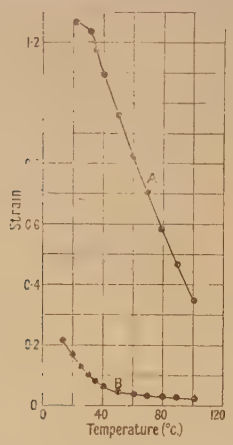


Figure 8. Strain–temperature curves for unloaded specimens.

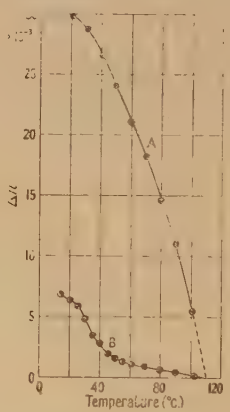


Figure 9. Birefringence–temperature curves for unloaded specimens.

In order to compare the theoretical expression with the experimental results, k has to be expressed in terms of the observed strain. This was done by assuming that there was no volume change in the material on stretching. Separate experiments have shown that this will involve an error of less than 1%. Figure 10 compares the expression with the experimental results at the lower temperatures, and it can be seen that the theoretical curve is of the same shape as the experimental, but the values are considerably lower. This is to be expected, since in obtaining the expression, it has been assumed that the limiting value is obtained at infinite strain, whereas for polyethylene, the molecular chains will pull the crystallites into complete alignment at a strain of about 4. The treatment is in any case approximate, since the birefringence due to the orientation of the amorphous material has not been considered.

The birefringence produced on stretching a network of molecular chains has been considered by Kuhn and Grün (1942) and Treloar (1947). The curve obtained from this treatment, assuming the same value for $n_2 - n_1$, is shown for comparison in Figure 10 as the 'theoretical curve for amorphous material'.

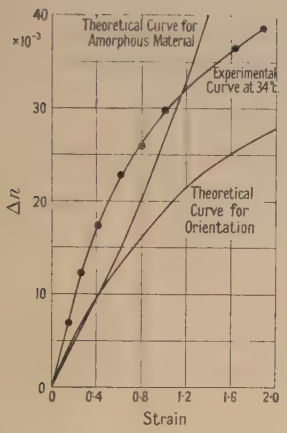


Figure 10.

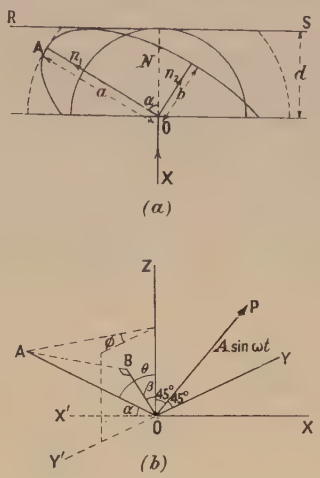


Figure 11.

It may be seen that it differs very little from the other theoretical curve in the linear region, but beyond this, the rate of rise increases with increasing strain, which is contrary to the behaviour of polyethylene. The Kuhn and Grün treatment also shows that the linear relation between stress and birefringence should hold over a greater range than that between strain and birefringence. For polyethylene, the reverse is found to be the case, and these differences must be attributed to its polycrystalline structure.

The marked decrease in gradient of the birefringence–strain curves for temperatures above 50° C. (Figure 5) may be attributed to reduction in percentage crystallinity of the material which has been observed by Bunn and Alcock (1945). Raine, Richards and Ryder (1945) deduced a relation between the percentage crystallinity and temperature from the differences in heat content of liquid and solid polyethylene and also from the heat of solution in a hydrocarbon. The resulting curve is similar to Figure 5. Measurements by Hunter and Oakes (1945) on the change of density of polyethylene with temperature are also in agreement with these results. Figure 7 indicates that crystal growth along the direction of stretching can still occur within 10° C. of the *softening point*, so that the complete structural disorientation associated with melting appears to take place more suddenly than the above discussion would suggest.

The strain–birefringence graphs (Figures 4 and 6) show that when the stress is reduced, the curves are retraced so that each value of the strain corresponds to a definite reversible orientation, and the *permanent set* found when the stress has been removed, is accompanied by a corresponding optical strain. There thus appears to be no evidence for slipping between the long chain molecules in the temperature range considered. It therefore seems unlikely that the behaviour of high polymers of this type can be adequately described in terms of simple

mechanical models of the spring and dashpot type, or in fact by any linear type of memory function.

ACKNOWLEDGMENTS

The authors wish to thank Mr. C. W. Bunn and Dr. L. R. G. Treloar for useful discussions, and Messrs. D. A. Gray and M. R. Davis for help with the experimental work.

REFERENCES

Bunn, C. W., 1941, *J. Sci. Instrum.*, **18**, 70; 1942, *Proc. Roy. Soc.* A, **180**, 40; 1949, *Fibre Science* (Manchester: Textile Institute), p. 144.
Bunn, C. W., and Alcock, T. E., 1945, *Trans. Faraday Soc.*, **41**, 317.
Hunter, E., and Oakes, W. G., 1945, *Trans. Faraday Soc.*, **41**, 49.
Kolsky, H., and Shearman, A. C., 1943, *Proc. Phys. Soc.*, **55**, 383.
Kuhn, W., and Grün, F., 1942, *Kolloidzschr.*, **101**, 248.
Raine, H. C., Richards, R. B., and Ryder, H., 1945, *Trans. Faraday Soc.*, **41**, 56.
Richards, R. B., 1945, *Trans. Faraday Soc.*, **41**, 127.
Treloar, L. R. G., 1947, *Trans. Faraday Soc.*, **43**, 277.

APPENDIX

Calculation of the Birefringence Produced by a System of Oriented Crystals

Consider first the propagation of light through a single uniaxial crystal. The light is incident normally, the axis OA of the crystal making an angle α with the direction of the light OX (Figure 11 (*a*)).

Let the ordinary and extraordinary refractive indices be n_1 and n_2 respectively. Then if the thickness of the crystal is d, the optical path length for the ordinary ray is $n_1 d$. To find the path length for the extraordinary ray, the ellipsoidal wave surface must be considered using Huyghens' principle.

From the geometry of the system it can be seen that

$$d^2 = a^2 \cos^2 \alpha + b^2 \sin^2 \alpha \qquad \ldots\ldots (1)$$

where a and b are the major and minor axes of the ellipsoid.

The extraordinary ray will have just traversed the crystal when RS is tangential to the ellipsoidal wave surface (Figure 11 (*a*)).

From equation (1), the time taken for the extraordinary ray to traverse the crystal will thus be given by

$$an_1/\mathbf{c} = bn_2/\mathbf{c} = dn_1 n_2/\mathbf{c}(n_1^2 \sin^2 \alpha + n_2^2 \cos^2 \alpha)^{1/2} \qquad \ldots\ldots (2)$$

where $\mathbf{c}$ is the velocity of light *in vacuo*.

The effective path length for the extraordinary ray will therefore be given by Nd where

$$N = n_1 n_2 (n_1^2 \sin^2 \alpha + n_2^2 \cos^2 \alpha)^{-1/2}. \qquad \ldots\ldots (3)$$

In Figure 11 (*b*), the crystal is shown in relation to three rectangular axes, OX, OY, OZ, the plane OAX making an angle β with the plane XOZ.

If light polarized with its electric vector along OP, which makes an angle of 45° with OZ, passes through the crystal, the phase difference between the components of the emergent light polarized along OZ and OY will give a measure of the birefringence.

The ordinary ray will be polarized with its electric vector perpendicular to the optic axis OA. The extraordinary ray is polarized at right angles to this in the direction OB lying in the YOZ plane.

If the incident light vector is $A \sin \omega t$, the components along and perpendicular to OB are $A \sin \omega t \cos(\beta + \pi/4)$ and $A \sin \omega t \sin(\beta + \pi/4)$.

After passing through the crystal the emergent vectors are

$A \sin [\omega t + (2\pi Nd)/\lambda] \cos (\beta + \pi/4)$ and $A \sin [\omega t + (2\pi n_1 d)/\lambda] \sin (\beta + \pi/4)$.

Resolving along OZ gives

$A \sin [\omega t + (2\pi Nd)/\lambda] \cos (\beta + \pi/4) \cos \beta + A \sin [\omega t + (2\pi n_1 d)/\lambda] \sin (\beta + \pi/4) \sin \beta$,

and along OY,

$-A \sin [\omega t + (2\pi Nd)/\lambda] \cos (\beta + \pi/4) \sin \beta + A \sin [\omega t + (2\pi n_1 d)/\lambda] \sin (\beta + \pi/4) \cos \beta$.

These can be written respectively as

$$Z_1 \sin [\omega t + \pi d(N + n_1)/\lambda + \delta_1] \quad \text{and} \quad Z_2 \sin [\omega t + \pi d(N + n_1)/\lambda + \delta_2],$$

where

$$2Z_1^2 = A^2[1 - \tfrac{1}{2}\sin 4\beta\{1 - \cos [2\pi d(N - n_1)/\lambda]\}] \quad \ldots\ldots (4)$$

$$2Z_2^2 = A^2[1 + \tfrac{1}{2}\sin 4\beta\{1 - \cos [2\pi d(N - n_1)/\lambda]\}] \quad \ldots\ldots (5)$$

$$\tan \delta_1 = \{\cos 2\beta - \sin 2\beta\} \tan \{\pi d(N - n_1)/\lambda\} \quad \ldots\ldots (6)$$

$$\tan \delta_2 = -\{\cos 2\beta + \sin 2\beta\} \tan \{\pi d(N - n_1)/\lambda\} \quad \ldots\ldots (7)$$

If $d(N - n_1)$ is small, $Z_1 = Z_2 = A/\sqrt{2}$,

$$\delta_1 = [\cos 2\beta - \sin 2\beta]\pi d(N - n_1)/\lambda, \qquad \delta_2 = -[\cos 2\beta + \sin 2\beta]\pi d(N - n_1)/\lambda,$$

so that the phase difference introduced is

$$\delta_1 - \delta_2 = [2\pi d(N - n_1) \cos 2\beta]/\lambda. \quad \ldots\ldots (8)$$

In considering the birefringence produced by the system of crystallites, the direction of stretch perpendicular to OX is an axis of symmetry, and it is convenient to change the coordinates. Taking OZ as the direction of stretch, let angle AOZ be θ and let the angle between the planes AOZ and ZOY be ϕ.

Substituting in equation (8) for α and β, N being given by equation (3), the phase difference becomes

$$\delta_1 - \delta_2 = \{(2\pi d)/\lambda\}\left\{\frac{n_2}{(1 + m \sin^2\theta \sin^2\phi)^{1/2}} - n_1\right\}\left\{\frac{2\cos^2\theta}{(1 - \sin^2\theta \sin^2\phi)} - 1\right\} \quad \ldots\ldots (9)$$

where $m = (n_2^2 - n_1^2)/n_1^2$. In the unoriented material, the number of crystals whose axes lie between θ' and $\theta' + \delta\theta'$, and ϕ' and $\phi' + \delta\phi'$ is given by $(M\delta\theta' \sin\theta' \delta\phi')/2\pi$, where M is the total number of crystals. If the crystals are rod-shaped and embedded in a matrix which is stretched, these crystals will then lie between θ and $\theta + \delta\theta$ and ϕ and $\phi + \delta\phi$, where $\phi = \phi'$ and $\tan\theta = k \tan\theta'$, k being the ratio between the width and length of a square of material after stretching.

The total phase difference for a large number of small crystals is then given by

$$\frac{2Mkd}{\lambda} \int_{\eta=0}^{\eta=\pi/2} \int_{\eta=0}^{\eta=\pi} \frac{\tan\theta \sec^2\theta}{(k^2 + \tan^2\theta)^{3/2}} \left\{\frac{n_2}{(1 + m \sin^2\theta \sin^2\phi)^{1/2}} - n_1\right\} \times \left\{\frac{2\cos^2\theta}{1 - \sin^2\theta \sin^2\phi} - 1\right\} d\theta\, d\phi.$$

This double integral may be evaluated by expanding $(1 + m \sin^2\theta \sin^2\phi)^{-1/2}$ as a series in powers of m.

Inserting the values of n_1 and n_2 for polyethylene (Bunn 1949), m becomes 0·059 and it is found that an error of less than 1% is introduced by ignoring powers of m higher than the first.

The value for the birefringence is then $\frac{n_2 - n_1}{2}\left\{\frac{3}{(1 - k^2)} - \frac{3k \cos^{-1} k}{(1 - k^2)^{3/2}} - 1\right\}$, taking $m = 2(n_2 - n_1)/n_1$.

The Friction and Mechanical Properties of Powders Bonded by Synthetic Resins

By R. C. PARKER and E. J. W. WHITTAKER
Messrs. Ferodo Limited, Chapel-en-le-Frith, Stockport

Communicated by E. N. da C. Andrade; *MS. received* 31*st July* 1950

ABSTRACT. The coefficient of friction and mechanical properties have been determined for a series of non-metallic materials formed by bonding mineral powders with resins. The manner in which the coefficient of friction varies with particle size has been studied, and a tentative explanation has been advanced involving certain geometrical parameters. A correlation was observed between the coefficient of friction and the hardness of the mineral powders expressed on Moh's scale. Marked correlations were also obtained between compression strength, cross-breaking strength and the reciprocal of the particle size for most of the materials; consideration is given to the various mechanical and thermodynamical factors involved.

§ 1. INTRODUCTION

There is no general agreement on the basic mechanism of friction. The asperity theory, first postulated by Coulomb (1785), is still adhered to by Bickerman (1948), while the cohesive theory, first suggested by Hardy (1936), is supported by the work of Bowden and collaborators (Bowden and Tabor 1939). The work of Bowden, in particular, emphasizes the complex nature of friction, and shows that the bulk properties of the solid are involved in addition to the surface properties. Schnurmann and Warlow-Davies (1941) suggest that the force of sliding friction may include an electrostatic component, and cite experimental evidence in support.

It must, however, be remembered that solid friction is met with under a wide range of conditions, so that no one mechanism is likely to have general application. For example, the experiments of Ernst and Merchant (1940) show that with very rough surfaces asperities contribute to friction, while recent work by Holm (1946) and by Parker and Hatch (1950) describes experiments in which the tangential component of friction is solely attributed to cohesive forces. The factors on which friction depends must be governed by the precise conditions prevailing at the instant of the experiment.

The explanations of friction now current owe their origin to Coulomb's law, although Parker (1949) has pointed out that this law has only very limited application. Other laws have, of course, been advanced from time to time, and each has added to the picture of friction.

A study of the literature on friction shows that attention has largely been confined to metallic materials, although a number of workers have used glass as one of the sliding members. It was therefore felt that a further study of non-metallic materials might open fresh ground and possibly lead to some new law which, though restricted in scope, might nevertheless add something to our knowledge of friction. A non-metallic material was sought that would enable control to be exercised over the surface and bulk properties. No one material or synthetic substance seemed suitable, so research was made into a compound material formed by bonding powders with a resin. By suitable choice of a powder a wide range of properties was covered.

Preliminary experiments showed that the mechanical properties of these composite materials were also sensitive to the variables considered in the friction experiments. These properties were therefore measured, and are also reported here in view of the possible relevance of bulk properties in frictional phenomena and also because they are of some intrinsic interest.

§ 2. PREPARATION OF MATERIALS

The minerals used in the investigation were calcite, fluorite, magnetite and silica, which were all readily available and which covered a wide range of hardness. The specimens were obtained by crushing the hand-picked minerals with the exception of silica which was obtained as a commercial powder. In all cases the powders were graded by sieving on a shaking machine, and subsequently freed from adhering dust by wet sieving on the appropriate standard sieve. Standard Institute of Mining and Metallurgy (IMM) sieves were used to isolate the following five fractions : $-20+40$; $-40+70$; $-70+120$; $-120+200$; and $-200+300$ IMM. The fraction -300 IMM to dust was discarded. The grade $-20+40$ IMM was not available for silica.

Two phenolic resins were used as bonding agents. The first (denoted as resin I) was available as a fine powder, and the second (resin II) as a solvent-free liquid. The powders were bonded into rectangular blocks measuring $2 \times \frac{1}{2} \times \frac{1}{2}$ in. The necessary weight of mineral powder was loaded into a die of internal dimensions $2 \times \frac{1}{2} \times 1\frac{1}{2}$ in. deep, and mixed with an excess of resin. This excess was removed by the application of pressure while the resin was in a liquid condition, this condition being attained in the case of resin I by heating the die and its contents to 100° C. in an electric oven. The pressure was applied to the stone–resin mixtures by means of a suitably designed close-fitting plunger. The pressure, which, incidentally, slightly compacted the powder, was limited to a value less than that which would cause crushing of the powder. Preliminary experiments showed that 220 lb in^2 did not crush calcite, and this value was standardized throughout the work, since the other minerals were harder and more resistant to crushing. After pressing, the resin component was cured by the application of heat.

Blocks were prepared from each of the minerals bonded with resin I, and from calcite and silica bonded with resin II. Six combinations of the materials are therefore considered.

It is to be expected that a material prepared by bonding a powder with a synthetic resin will possess the greatest mechanical strength if the proportions of the powder and bonding agent are such that the latter just fills the voids in the former. This expectation was verified early in this work, and every effort was therefore made to ensure that the voids were filled as completely as possible.

§ 3. MEASUREMENT OF FRICTION AND MECHANICAL PROPERTIES

A side view of the apparatus used for measuring the kinetic coefficient of friction is shown in Figure 1. A torque arm A carries two test pieces which bear on a circular mild steel plate B. The torque arm is pivoted in the plane of the bearing surface, and the test pieces are held in screw-type clamps near either end of the arm. The plate B is directly connected to a 3 H.P. motor and is driven by a 'Ward-Leonard' unit. The normal load is applied by hanging lead weights on the lever arm, and the thrust is transmitted by three balls disposed round the

centre of A by means of a hemispherically ended connecting link. The arm is constrained by two light cords that pass from each end of the arm, over pulleys, to tungsten-alloy weights which hang in containers filled with mercury. In the plan view of the apparatus, Figure 2, may be seen two brass guides D which ensure that the cord acts at a constant radius. On the steel rod E is fixed a pen that

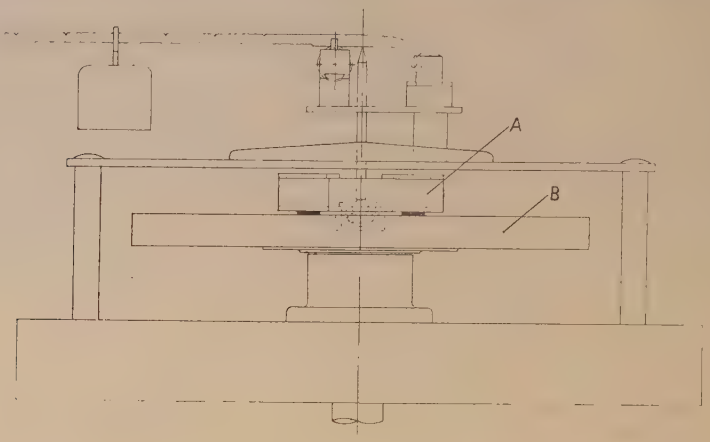


Figure 1.

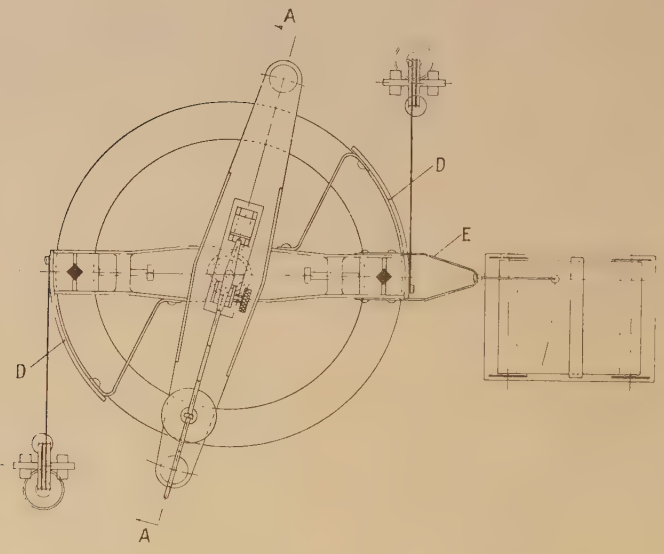


Figure 2.

records the deflection of the torque arm on a moving chart. The chart is driven by a constant speed motor that can be geared to give selected speeds between 3 in. per min. and 3 in. per hr.

In the original design of the apparatus the tungsten-alloy weights were counterbalanced so that, in the absence of applied torque, the upper surface of the mercury was level with the top surface of the weight. The weights themselves were generated from an exponential curve and so gave linear sensitivity. However, for this work a greater range of torque was required, and this was most easily

provided for by hanging additional weights on the cord. In this way a range could be selected to cover a particular series of experiments.

The experimental procedure adopted was to run the machine up to speed with the pressure on, wait until the friction settled down, and then break the circuit of the 3 H.P. motor. The friction between the test pieces and plate was then recorded through the period of deceleration of the plate and armature. The value of friction recorded as the plate came to rest is described in what follows as the friction at zero speed.

Friction records were taken using the $\frac{1}{2} \times \frac{1}{2}$ in. face of the test pieces as the bearing surface under three sets of conditions.: an initial linear velocity of 50 ft/sec. with total normal loads of 25 lb. and 5·1 lb., and a linear velocity of 25 ft/sec. with a total normal load of 5·1 lb.

In many cases the coefficient of friction could not be determined at zero speed and at a normal load of 25 lb. since the friction readings involved overlapping two ranges of the recording mechanism, and also because the change in friction with time was so rapid that inertia effects of the recording mechanism made the results unreliable.

The cross-breaking strength was also measured on $2 \times \frac{1}{2} \times \frac{1}{2}$ in. rectangular blocks. The specimens were laid symmetrically on parallel V-shaped edges 1 in. apart, and a load applied to the opposite face by a third V-shaped edge, parallel to, and midway between, the supporting blocks. The edges had a radius of $\frac{1}{16}$ in. and were 1 in. in length. The compression strength was measured on $\frac{1}{2}$ in. cubes cut from the rectangular blocks.

§ 4. FRICTION

(*a*) *The Effect of Particle Size*

The effect of particle size on the coefficient of friction of the various materials is shown in Figures 3 (*a*)–(*d*). The values of friction plotted are the means from three to four repeated experiments on each particle size for each material, and the result from each of these experiments was the mean of six consecutive tests. The particle size shown on the abscissa is the size of the opening in the larger of the two sieves defining a sieve grade. This was found to be sensibly equal to the diameter of the equivalent sphere for calcite particles, which was obtained from the density of calcite and the weight of counted numbers of particles.

A study of Figures 3 (*a*)–(*d*) shows that only in the case of silica has particle size a consistently marked effect at finite velocities. In general, the coefficient of friction tends to rise to a limiting value with diminishing particle size with a maximum gradient in the region of the larger particle sizes considered. This type of variation is most persistent in the case of silica, with respect both to changes in testing conditions and to changes of bonding resin. Other minerals showed, under certain conditions, quite divergent types of variation, or none at all.

In the accompanying Table the crosses indicate those cases for which statistically significant variations in the coefficient of friction with variation of the particle size were found either by application of Student's t-test or by calculation of correlation coefficients.

No straightforward explanation of the effect of particle size on the frictional properties is possible in the present state of knowledge of the fundamental nature of friction between a metal and a non-metal. However, the variation of the coefficient of friction with pa · icle size suggests an effect occurring at the boundaries

		25 lb. (50 ft/sec.)	5·1 lb. (50 ft/sec.)	5·1 lb. (25 ft/sec.)	5·1 lb. Zero speed
Resin I	Calcite	—	—	×	×
	Fluorite	×	—	×	×
	Magnetite	×	—	×	—
	Silica	×	×	×	×
Resin II	Calcite	×*	—	—	—
	Silica	×	×	—	×

* This gave a significant negative correlation coefficient between friction and reciprocal particle size, whereas all the others recorded were positive.

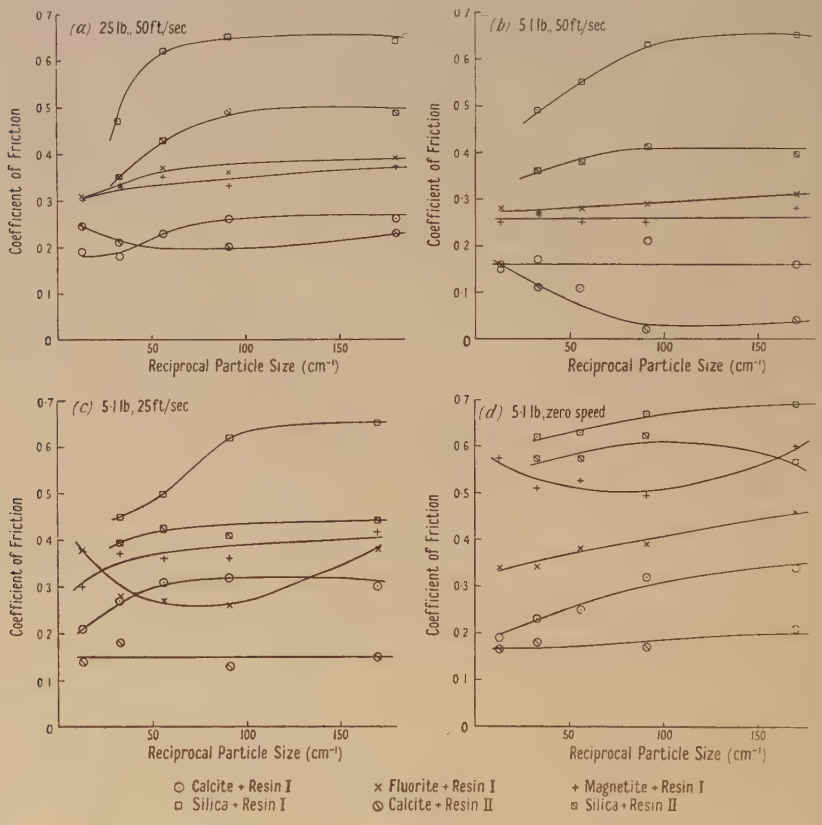


Figure 3.

between the mineral particles and the bonding resin since the most frequent type of variation involves an initial rise of friction with decreasing particle size. This phenomenon may be connected with one of the two geometrical parameters of a plane cutting a two-phase body which have inverse proportionality to the particle size of the disperse phase when the proportions of the two phases and the particle shape of the disperse phase are kept constant. These are the total inter-phase perimeter per unit area, and the number of times such a boundary is traversed by a unit line lying in the surface. It is therefore suggested that it is possible that some action occurs at the inter-phase boundary which affects the frictional behaviour of the material, and that this is likely to arise from the effect of the wear products of one phase on the frictional behaviour of the other.

(*b*) *The Effect of Mineral Hardness*

The coefficient of friction between the resin–mineral blocks and mild steel is in general found to be correlated with the hardness of the mineral under all the conditions studied. This may be seen to some extent in Figures 3 (*a*) – (*d*).

The full value of these results can only be obtained by isolating the effects of hardness from the effects of particle size. This can be done by first expressing the coefficient of friction of each material for a given particle size as a fraction of the average of the values for all the minerals at that particle size. The relative values thus obtained for each mineral are then plotted against the hardness of the mineral on Moh's scale. This procedure was applied separately to the materials bonded with the two different resins. The resulting plot for the values corresponding to zero speed at the end of deceleration tests at a load of 5·1 lb. is shown in Figure 4.

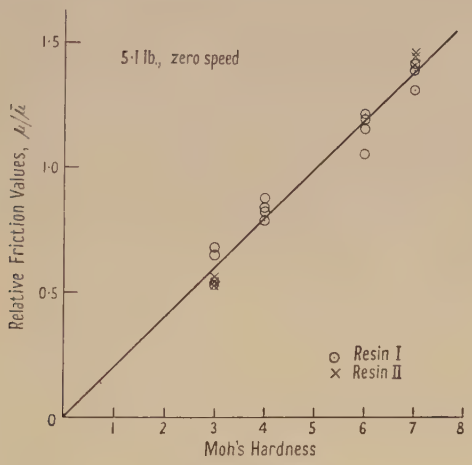


Figure 4.

The linearity of the relationship, the fact that the line appears to pass through the origin, and the fact that the results from both resin I and II lie on the same line, are all remarkable. A similar relationship was shown by results measured at the same pressure at the finite test velocities.

We may therefore draw the following conclusions with regard to the relation between the hardness of a mineral and the coefficient of friction against mild steel which is given by resin-bonded blocks prepared from it.

(i) Hardness is the most important characteristic of the mineral in determining the frictional properties of such blocks.

(ii) Under kinetic conditions some other properties of the mineral do play a small part, although hardness is the major factor, but at zero speed (which is to be distinguished from static friction as normally defined) these complicating effects do not occur appreciably.

(iii) The effect of hardness on friction is not sensitive to the properties of the bonding resin (over the range of these experiments) or to the particle size of the mineral powder.

(iv) Under the simplified conditions of zero speed, and when the particle size of the mineral and the properties of the bonding resin are held constant, the friction is directly proportional to the Moh's hardness of the mineral.

These conclusions may be symbolized mathematically as follows. The coefficient of friction μ is a function of the mineral hardness H, the properties R of the resin, the particle size d of the mineral powder, and probably other unknown parameters x which have not been specifically isolated, but which are almost independent of the nature of the mineral at low speed and pressure. Thus we may put $\mu = \mu(H, R, d, x)$. Then we have shown that

$$\frac{\mu(H_1, R_1, d_1, x_1)}{\mu(H_2, R_1, d_1, x_1)} = \frac{H_1}{H_2},$$

or, expressed differentially, $\left(\frac{\partial \mu}{\partial H}\right)_{R,d,x} = \frac{\mu}{H}$.

Although many workers have studied the relation between wear and hardness, there has hitherto been no determined effort to relate friction with hardness. That a relation exists for the material studied here is, therefore, surprising, and particularly so in view of the empirical nature of Moh's scale.

The reason for this relation cannot be advanced at this stage. The mechanism of scratching which it suggests can offer no direct explanation, since Whittaker (1947) has shown that, for the type of material considered here, that portion of abrasion which actually results in wear involves a negligible fraction of the frictional work. It may well be, however, that the degree of abrasion affects the friction by removal of the contaminating films to which, under normal conditions, the frictional coefficient is so sensitive.

§ 5. MECHANICAL STRENGTH

The resin-bonded mineral powders were all tested to fracture in both compression and cross-breaking tests as described in § 2 above, with the exception that the materials prepared from magnetite with resin I and from silica with resin II were not tested in compression. The results plotted against reciprocal particle size are shown in Figures 5 and 6. The values plotted are the means from five or six repeat experiments, except for magnetite with resin I and silica with resin II,

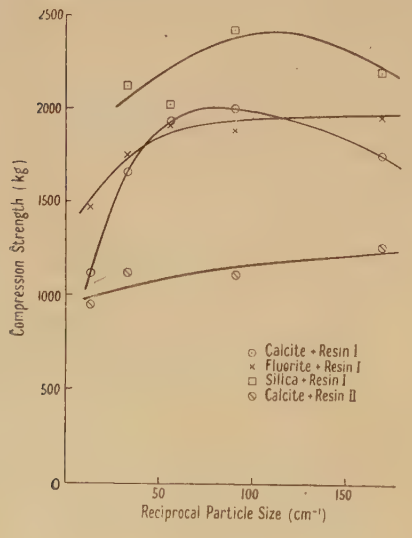


Figure 5.

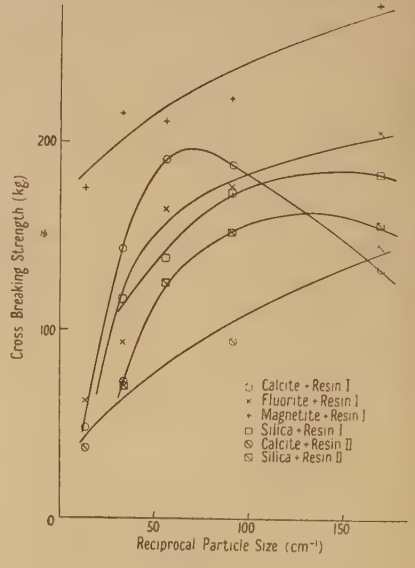


Figure 6.

for which only two results were obtained for each material. Statistical analysis of the whole body of data has shown statistically significant correlation coefficients between both types of strength measurements and the reciprocal particle size for most of the materials, and the general similarity of all the curves is good evidence that the relationship is real in all cases.

The compression and cross-breaking strength results differ in that the range of variation of the former due to variations in the nature of the materials is greater than is produced by variation of the particle size of the mineral, whereas the reverse is true of the cross-breaking strength results. The two properties resemble each other, however, in the general form of the relationship with the particle size of the mineral. Calcite with resin I gives curves with a definite maximum; silica with resin I probably also shows a maximum, and as the slope of the curve for the other materials decreases with decrease in particle size it may well be that a maximum occurs for particles smaller than those studied.

Microscopic examination of the resin-bonded materials shows that the resin wets the mineral powder. On the other hand there is no doubt that fractured surfaces always contain a large proportion of exposed surfaces of mineral particles, i.e. fracture occurs more readily at the phase boundary than within either component. Two explanations are possible : (i) The wetting of the mineral powder by the resin may only imply that the attractive forces between the mineral and the resin are greater than those within the resin while the latter is in the liquid state, and that the internal forces in the resin become the greater after curing. Such a change as this would be likely only if a considerable contribution to the strength of the resin were made by primary valency forces, and this is improbable. (ii) The mineral-to-resin adhesion may be greater than the internal cohesion of the resin in accordance with the evidence of wetting, but the stress at the phase boundary may be higher than elsewhere owing to the difference in elastic moduli of the two components. This explanation appears highly probable. Experimental evidence in its favour may be adduced from some measurements which have been made on specimens of the cured resins themselves, which indicated that the effect of the mineral filler can either increase or decrease the strength of the material as compared with the unfilled resin. The view outlined in (i) above could not explain the existence of filled material with higher mechanical strength than that of the resin, whereas (ii) would admit of such a result if the ratio of the stress concentration at the interface to that in the resin exceeded the ratio of the strength of the resin–mineral and resin–resin bonds. It is therefore suggested that the factors involved in the variation of mechanical strength with particle size will be of two types : the thermodynamic factors associated with the adhesion energy between the resin and the mineral and the mechanical factors affecting the stress distribution. These may be analysed further as follows :

1. *Thermodynamic factors.* The adhesion energy at the mineral–resin interface will depend, for given materials, on the extent of the interface and on its curvature. If we neglect the effect of the curvature of the interface the change in free energy per unit volume involved in separating the resin from the mineral particles will be given by $\Delta F = S(\gamma_1 + \gamma_2 - \gamma_3)$, where γ_1, γ_2, γ_3 are the surface tensions of the resin–air, mineral–air, and mineral–resin interfaces respectively, and S is the total area of interface per unit volume.

If the total volume of the particles per unit volume is constant (which was nearly true in this work), S is then directly proportional to the specific surface of the

powder for a given material, and so inversely proportional to particle size for geometrically similar particles. The adhesion energy is therefore also inversely proportional to the particle size.

Allowance for the effect of the curvature of the surface of the particles does not modify this conclusion. If the particles are assumed to be spherical, of radius r, the free energy change per unit volume becomes

$$\Delta F = S(\gamma_1+\gamma_2-\gamma_3) - \frac{2}{r}\{\gamma_1(1-v)-\gamma_2 v-\gamma_3(1-2v\},$$

where v is the fraction of the volume occupied by the particles, so that the inverse proportionality to particle size still holds true. It is clear, however, that this analysis does not establish the sign of the proportionality constant in the absence of information on the magnitudes of the surface tensions, and it must be expected in any case that the effect of this thermodynamic factor will be considerably modified by the mechanical factors.

2. *Mechanical factors.* This heading covers a number of factors in connection with which the effect of particle size cannot be predicted theoretically. It is clear, however, that they must be expected to modify the strict applicability of the thermodynamic analysis: (*a*) residual stress, particularly at the inter-phase boundaries, resulting from volume changes in the resin on curing; (*b*) residual stress, again particularly at the interfaces, resulting from differential thermal contraction of the components on cooling from the curing temperature; (*c*) localization of applied stress resulting from the different elastic moduli of the components.

It is therefore concluded that the dependence of the mechanical properties on the size of the mineral particles can be explained qualitatively, but that a quantitative explanation, even for particles of a simple shape, would require much more information about the stress distribution in such inhomogeneous materials and about the surface tensions of the components than is available at present.

ACKNOWLEDGMENTS

The authors wish to thank Mr. D. Hatch, who carried out most of the experimental work described in this paper. Special thanks are also due to Professor E. N. da C. Andrade for reading the manuscript and for making valuable suggestions regarding its manner of presentation. Finally, the authors are indebted to Mr. William Smith, Chairman of Ferodo Limited, for his kind permission to publish this paper.

REFERENCES

Bickerman, J. J., 1948, *Lubricating Engng.*, **4**, 208.
Bowden, F. P., and Tabor, D., 1939, *Proc. Roy. Soc.* A, **169**, 391; 1945, *A. R. Chem. Soc.*, **42**, 20.
Coulomb, A., 1785, *Mém. Acad. Roy. Sci.*, p. 161.
Ernst, H., and Merchant, M. E., 1940, *Conference on Friction and Surface Finish, M.I.T.*, p. 76.
Hardy, W. B., 1936, *Collected Scientific Papers of the Cambridge University Press.*
Holm, R., 1946, *Electric Contacts* (Stockholm: Almquist and Wiksells).
Parker, R. C., 1949, *Engineering*, **167**, 193, 217.
Parker, R. C., and Hatch, D., 1950, *Proc. Phys. Soc.* B, **63**, 185.
Schnurmann, R., and Warlow-Davies, E., 1942, *Proc. Phys. Soc.*, **54**, 14.
Whittaker, E. J. W., 1947, *Nature, Lond.*, **159**, 541.

Magnetic Susceptibility and Anisotropy of Mica

By J. T. KENDALL* and D. YEO†
Research Department, Metropolitan-Vickers Electrical Co. Ltd.

MS. received 8*th September* 1949, *and in amended form* 24*th July* 1950

ABSTRACT. The magnetic susceptibility and anisotropy of natural muscovite and synthetic fluorphlogopite have been measured. These minerals contain small quantities of iron which cause paramagnetic and ferromagnetic effects. To a fair approximation the mean susceptibility is found to vary linearly with the total iron content, and the paramagnetic anisotropy to be proportional to the ferrous iron content. At low field strengths ferromagnetic impurities (probably sub-microscopic inclusions of magnetite) cause anomalous results. It is shown that such inclusions are present in all the micas examined, and account for the apparent paramagnetism of synthetic mica. The inclusions are probably orientated with their magnetic axes in a preferred direction relative to the crystallographic axes of the mica.

§ 1. INTRODUCTION

Comparatively large crystals of synthetic mica were produced in Germany during the war by Middel (1946) and Eitel and Dietzel (1946). It is reported by the latter that their synthetic mica was paramagnetic, in spite of its small iron content. Synthetic mica has also been made here (Kendall and Spraggon 1947), and we have reported briefly on its magnetic susceptibility (Kendall and Yeo 1948). We found that synthetic mica prepared on a large scale by slow cooling was diamagnetic, as expected, but that when prepared on a small scale it exhibited a weak ferromagnetism, probably due to the presence of unabsorbed sub-microscopic inclusions of magnetite as impurity. In all cases the iron content was less than 0·05%. This work has now been extended to include a study of the magnetic anisotropy. The following types of mica have been studied : (i) natural muscovite from Madras, (ii) natural muscovite from Calcutta, (iii) synthetic fluorphlogopite made by Middel, (iv) synthetic fluorphlogopite made by Kendall and Spraggon.

According to the modern theories of magnetism mica should be diamagnetic. In fact a rough calculation based on the additivity of specific ionic susceptibilities indicates that a mineral corresponding to fluorphlogopite ($KMg_3AlSi_3O_{10}F_2$) should have a specific diamagnetic susceptibility of $-0{\cdot}35\times10^{-6}$ c.g.s. units. The incorporation of sufficient iron in the crystal lattice will, however, make the mineral paramagnetic. Using the usual Bohr magneton value for iron of $p_B=5$, and calculating the paramagnetic susceptibility of the iron from the relation $p_B=2{\cdot}84(\chi_i T)^{1/2}$, where χ_i is the susceptibility per gram-ion and T is the absolute temperature, one finds that 0·2% of iron corresponds to a paramagnetic susceptibility of just $0{\cdot}35\times10^{-6}$. Synthetic mica with greater iron content should therefore be paramagnetic, but with less iron should be diamagnetic.

Wilson (1920) measured the susceptibilities of various clear and spotted varieties of muscovite, and found that in all cases the crystals had anisotropic magnetic properties, the susceptibility χ parallel to the cleavage plane being always greater than the susceptibility $\chi_\perp$ perpendicular to it. He also showed that inclusions caused enormous variation in the magnetic properties, and tentatively suggested that the anisotropy might be due either to a preferred orientation of the

* Now Leverhulme Research Fellow, Imperial College, London.
† Now with the Colonial Insecticides Research Project, Tanganyika.

inclusions or to an intrinsic anisotropy of the mica itself. However, no attempt was made to distinguish between the two possibilities.

Nilakantan (1938) measured the magnetic properties of various clear specimens of biotite and one each of phlogopite and muscovite. The specimens were carefully chosen to be free from any microscopic inclusions; nevertheless it is likely that they contained inclusions of sub-microscopic size. He found that the mean susceptibility was directly proportional to the total amount of iron present, while the anisotropy $\chi_{\parallel} - \chi_{\perp}$ was proportional to the amount of *ferrous* iron. For the biotites and phlogopite the anisotropy in the cleavage plane was negligibly small but for muscovite was measurable.

§ 2. MEASUREMENT OF ANISOTROPY

The general principles of magnetic anisotropy have been described fully elsewhere (Jackson 1924, Krishnan *et al.* 1933, Lonsdale 1937). Briefly, the magnetic susceptibility of a crystalline substance may be described by a magnetic ellipsoid. If the susceptibilities are measured for the three principal axes of this ellipsoid, and the positions of these axes fixed relative to the crystal axes, the magnetic properties of the substance are completely described.

Mica has, strictly, a monoclinic structure, but we have found it sufficiently accurate to use a magnetic ellipsoid with χ_1 perpendicular to the cleavage plane (thus corresponding to the c crystallographic axis) and χ_2 and χ_3 in the cleavage plane (thus corresponding respectively to the a and b crystallographic axes). Note that $\chi_2 > \chi_3 > \chi_1$.

A method of measuring the anisotropy was originally devised by Krishnan and his co-workers (1933). A crystal is freely suspended in a uniform horizontal magnetic field, and if the principal susceptibilities in the horizontal plane are χ_x and χ_y ($\chi_x > \chi_y$), we have

$$\text{Torque acting on crystal} = (\chi_x - \chi_y)\tfrac{1}{2}H^2 m \sin 2\alpha,$$

where H = magnetic field strength, m = mass of crystal, α = angle between χ_x and the direction of the field.

The torque is thus a maximum when $\alpha = \pi/4$. Suppose that the crystal is originally set with χ_x parallel to the field, then if the suspension is slowly twisted through an angle θ, the crystal will correspondingly rotate through a smaller angle α until $\alpha = \pi/4$, at which point the crystal will suddenly spin round. The position of the torsion head for $\alpha = \pi/4$ can thus be clearly fixed, and at this position we have

$$C\theta = (\chi_x - \chi_y)\tfrac{1}{2}H^2 m,$$

where C = torsional constant of the suspension, θ = angular rotation of the suspension (less $\pi/4$). From this $\chi_x - \chi_y$ may be calculated. Thus, using suitable crystal orientations $\chi_2 - \chi_1$, $\chi_3 - \chi_1$ and $\chi_2 - \chi_3$ may be found. The method is extremely sensitive, especially if the torsional constant of the suspension is such that it must be twisted through ten or twenty complete revolutions in order to move the crystal through an angle of $\pi/4$ in the magnetic field.

A fine quartz suspension (about 30 cm. long and 10–20 μ diameter) was attached at one end to a torsion head consisting of a pointer and a circular scale. For ease of working, a thin glass rod (5 mm. long and 0·2 mm. wide) was attached to the other end. The mica crystal was then attached to the free end of the rod, flake shellac, which had been shown to be free from ferromagnetism, being used for all joints. Cleavage flakes of mica were used, cut into the shape of a circle, the final specimens

being about 1 cm. diameter and 0·03 gm. weight. The torsional constant of the suspension was about 6×10^{-5} gm. cm^2 sec^{-2}. This may be measured by noting the time of rotation t when a thin glass disc of known moment of inertia I is attached to the fibre ($t = 2\pi(I/C)^{1/2}$). The field was produced by an electromagnet having plane parallel pole pieces 4 in. in diameter.

With the cleavage plane horizontal the specimen is symmetrical, and no correction due to anisotropy of shape and slight inhomogeneity of field is necessary; $\chi_2 - \chi_3$ is then measured directly, the directions of χ_2 and χ_3 being fixed relative to some reference line scratched on the cleavage surface. (The direction of χ_2 can easily be found since at zero *torque* the crystal will set with this direction parallel to the field. It is found to correspond with the *a* crystal axis.) With the cleavage plane vertical, measurements are carried out in two different positions, either χ_2 or χ_3 being vertical, thus giving measurements of $\chi_3 - \chi_1$ and $\chi_2 - \chi_1$. In these cases there might be a small correction necessary due to anisotropy of shape, but we found that suspension of the mica in a susceptibility bath of the same specific susceptibility as the mica did not affect the results, so that within our limits of experimental error the correction is negligible, the magnetic field being sufficiently homogeneous to avoid the effect.

Table 1. Magnetic Susceptibility and Anisotropy

Specimen		$\chi_2-\chi_1$ $\times 10^8$	$\chi_3-\chi_1$ $\times 10^8$	$\chi_2-\chi_3$ $\times 10^8$	χ_{mean} $\times 10^6$	% Fe^{++}	% Fe (total)
Muscovite (Calcutta)	a			9·3			
	b			9·3			
	c			9·3	+3·1	0·44	0·93
	d	52	42	6·9			
	e	36	30	6·4			
Muscovite (Madras Green)	f			5·7			
	g	34	30	3·8			
	h	33	30	3·1	+7·5	0·31	2·77
	i			3·0			
	j			3·7			
Fluorphlogopite (Middel)*	k	1·3	1·3	0·4	−0·24		0·05
Fluorphlogopite (Kendall and Spraggon)†	l				+1·6 (approx.)		0·04

* Large-scale preparation. † Small-scale preparation.

The experimental results obtained for five different specimens each of Calcutta and Madras muscovite and one of synthetic fluorphlogopite are shown in Table 1. The variation of the angular torque with magnetic field strength is shown in Figures 1–5 for the same specimens. The slope of these lines gives the anisotropy. It is noteworthy that Krishnan's method, when extended to include measurements at several field strengths, gives an excellent picture of any anomalous effects due to ferromagnetism. One would expect the lines to be straight and to pass through the origin, but in general they are curved and do not extrapolate through the origin. It is assumed that any anomalous ferromagnetic effects become less important at higher fields, and the results quoted in the Table are therefore taken from the upper straight-line portions of the curves.

§ 3. MEASUREMENT OF SUSCEPTIBILITY

The mean susceptibilities of optically clear natural muscovite and of fluorphlogopite prepared by us were measured by the Gouy method (i.e. by measuring the change in weight of a powdered sample due to a non-uniform magnetic field). A comparative method was employed, since it is difficult to determine field strengths and dimensions of the glass specimen tubes sufficiently accurately.

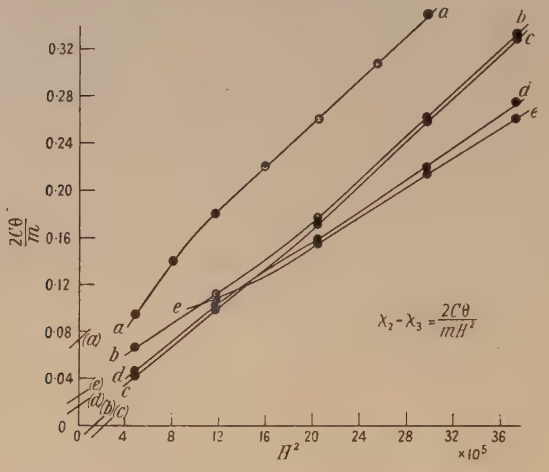


Figure 1. Anisotropy of muscovite (Calcutta), $\chi_2-\chi_3$.

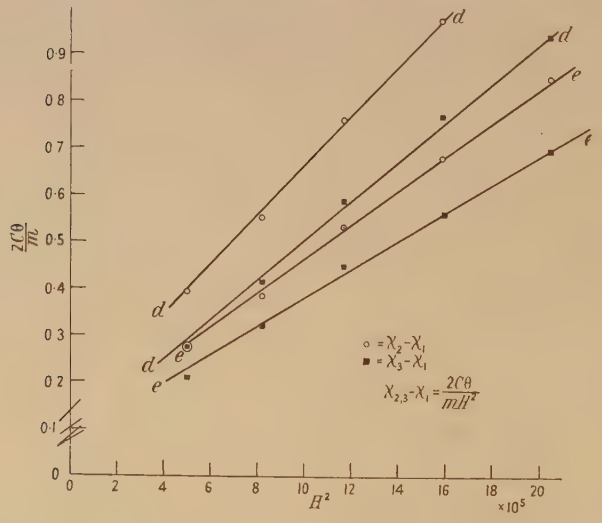


Figure 2. Anisotropy of muscovite (Calcutta), $\chi_2-\chi_1$ and $\chi_1-\chi_3$.

Water was used as the standard, its specific susceptibility being taken as $-0{\cdot}72\times10^{-6}$ c.g.s. units. Field strengths of 8,000 to 9,000 oersteds were used, no attempt being made to detect any variation of susceptibility with field strength, since the anisotropy measurements had shown that at these high fields ferromagnetic effects were absent, and our apparatus was not sensitive enough to warrant any study of the variation at low fields.

The mean specific susceptibility was calculated from the equation

$$\chi_M=\frac{1}{m}\left(\frac{K_a m}{\rho_s}\left(\chi_w W-\frac{K_a W}{\rho_w}\right)\left(\frac{F_s-F_a}{F_w-F_a}\right)\right),$$

where χ_M = mean specific susceptibility of the specimen, χ_w = specific susceptibility of water, K_a = volume susceptibility of air, ρ_s = density of specimen, ρ_w = density of water, m = mass of specimen in specimen tube, W = mass of water in same tube, F_s, F_w and F_a = changes in weight, with application of field, of the (tube and specimen), (tube and water) and (tube and air) respectively.

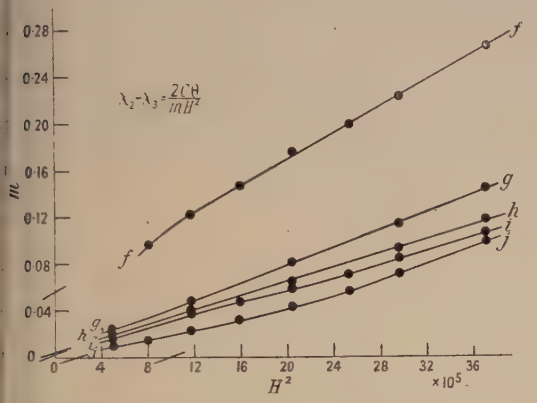


Figure 3. Anisotropy, muscovite (Madras Green), $\chi_2-\chi_3$.

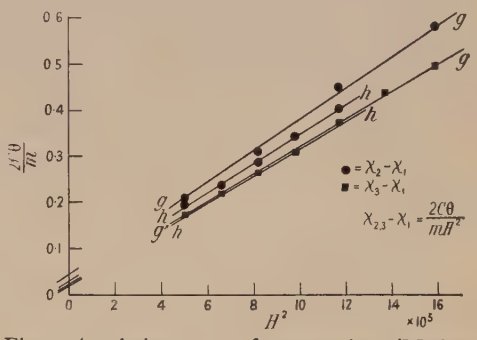


Figure 4. Anisotropy of muscovite (Madras Green), $\chi_2-\chi_1$ and $\chi_3-\chi_1$.

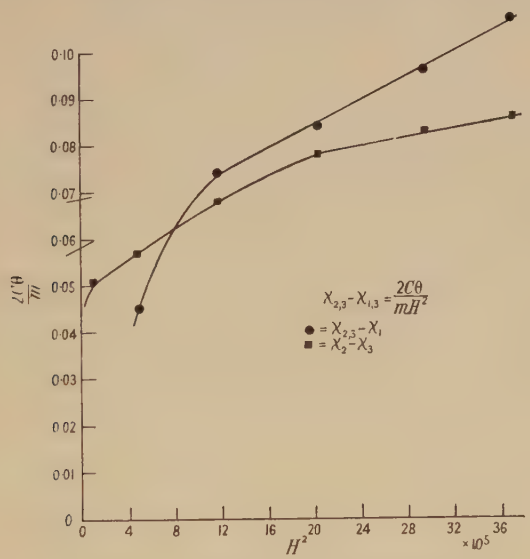


Figure 5. Anisotropy of fluorphlogopite.

The quantity of Middel's synthetic mica available was insufficient to use the Gouy method directly, and so a method due to Rabi (1927) was used as follows. A crystal was attached to a fine quartz fibre 50 cm. long, and suspended in a non-uniform field between conical pole pieces of an electromagnet. The crystal was surrounded by a solution the susceptibility of which could be varied, aqueous solutions of manganese chloride and potassium iodide being used. When the field was switched on the crystal moved, the direction of movement and rotation of the specimen depending on the magnitude and positions of the principal magnetic susceptibilities. The susceptibility of the solution was varied until there was no

movement at right angles to the axis of the magnet. The volume susceptibility of the specimen along the direction of the field was then the same as the volume susceptibility of the solution, which could be measured by the Gouy method as before. The experimental results are given in Table 1.

§ 4. CHEMICAL ANALYSIS OF SPECIMENS

The individual specimens were too small to allow an accurate estimation of their iron contents, and for the natural muscovites analyses were therefore carried out on the five specimens of each type taken together. The results are shown in the Tables. Standard analytical procedures were used for estimating the percentages of ferrous and total iron. In the case of the synthetic micas only the total iron could be estimated, the amount of ferrous iron being negligible.

§ 5. DISCUSSION OF RESULTS

Natural Muscovite

The results for the natural mica confirm those of Nilakantan (1938), namely that for clear specimens, free of inclusions when viewed under the microcope, the mean susceptibility varies linearly with the total iron content and the anisotropy $\chi_{\parallel} - \chi_{\perp}$ is proportional to the ferrous iron content. Thus the Madras muscovite has a larger total iron content than the Calcutta muscovite and a larger mean susceptibility, but a smaller ferrous iron content and a smaller anisotropy. This is in accord with Van Vleck's theory of crystalline paramagnetism, in which the ferrous, but not the ferric, ion can contribute to the anisotropy.

The anisotropy measurements vary somewhat from one crystal to another, as is only to be expected, since the iron content is unlikely to be very evenly distributed. A noticeable feature of Figures 1–4 is the curvature of the relation Torque/Field^2 at low fields, and this immediately suggests the presence of sub-microscopic ferromagnetic inclusions in the apparently clear mica, since diamagnetic and paramagnetic susceptibilities should be independent of field strength.

The effect of the ferromagnetic inclusions on the measured torque will depend on their anisotropy and upon the degree of orientation of their magnetic axes with respect to the crystallographic axes of the mica. At low magnetic fields the relation Torque/Field^2 will generally be curved; but at sufficiently high fields the ferromagnetic inclusions will become magnetically saturated—and if they are isotropic the additional torque due to them would disappear. The specimen would then behave as a normal anisotropic paramagnetic crystal and the relation Torque/Field^2 would be a straight line which, if extrapolated to zero field, would pass through the origin. The actual curves do in fact become straight at high fields—due to ferromagnetic saturation—but the extrapolated line does not in general pass through the origin. It must therefore be assumed that the ferromagnetic inclusions have anisotropic properties. The inclusions thus give a comparatively large torque at low fields, reaching a value at saturation which varies only slightly with the field. The upper straight-line portion of the curve, if extrapolated to zero field, does not then pass through the origin, but is displaced positively or negatively along the torque axis by an amount depending on the magnitude of the ferromagnetic anisotropy and the orientation of the resultant mean ferromagnetic axes relative to the paramagnetic axes. Magnetite, which has marked anisotropic properties (Weiss 1905), seems most likely to be responsible for the ferromagnetic behaviour. Unless the number of magnetite inclusions was

small, which seems unlikely, they must have been constrained by the crystal structure or crystalline electrostatic field to have a preferred orientation, otherwise no resultant anisotropy would have been detected.

Table 2. Ferromagnetic Anisotropy

Specimen	Extrapolated torque at zero field $(\chi_2\chi_1)$	$(\chi_3\chi_1)$	$(\chi_2\chi_3)$	% Fe^{++}	% Fe (total)
a			+0·073		
b			−0·014		
c			−0·020	0·44	0·93
d	+0·135	+0·085	+0·013		
e	+0·100	+0·080	+0·025		
f			+0·057		
g	+0·045	+0·025	+0·002		
h	+0·030	+0·025	+0·001	0·31	2·77
i			−0·007		
j			−0·038		
k	+0·058	+0·058	+0·068		0·05

The ferromagnetic anisotropies of the various specimens are proportional to the extrapolated values of the torque at zero field, and these are given in Table 2. It is seen that the values in the $(\chi_2\chi_3)$ plane are not at all constant, and, indeed, may be positive or negative, i.e. there is no orientation of the ferromagnetic axes in the cleavage plane, while in the planes at right angles to the cleavage some orientation is indicated.

Synthetic Mica

The synthetic fluorphlogopite made on a large scale by Middel (1946) is diamagnetic. Calculation based on the iron content (0·05%) and the chemical formula $KMg_3AlSi_3O_{10}F_2$ indicates a specific susceptibility of $-0{\cdot}25 \times 10^{-6}$ c.g.s. units. This corresponds closely with the mean experimental value given in Table 1. (Actually six specimens were examined, their susceptibilities being $-0{\cdot}17$, $-0{\cdot}29$, $-0{\cdot}19$, $-0{\cdot}28$, $-0{\cdot}26 \times 10^{-6}$, indicating that the small iron content was somewhat unevenly distributed.) The bulk of the iron was therefore incorporated in the crystal lattice.

However, the anisotropy measurements indicated that some iron was present in a ferromagnetic form, and reference to Figure 5 shows that at the fields used most of the torque is due to ferromagnetism. Only a very tiny amount of iron is needed to produce this effect.

The synthetic mica made here on a small scale was at first taken to be paramagnetic. However, the value given in Table 1 has little significance since the susceptibility varied considerably with the field strength (decreasing with increasing field), and it was also found that some of the iron could be removed by prolonged grinding and washing in dilute sulphuric acid. It seems likely, therefore, that a considerable portion of the iron was present as magnetite inclusions and was not incorporated in the crystal lattice. Suitable large crystals were not available, but very rough anisotropy measurements indicated a behaviour similar to that shown in Figure 5, but with values of torque about ten times larger.

ACKNOWLEDGMENTS

The authors wish to thank Dr. V. Middel for some of his synthetic mica (which found its way into their hands by a devious route) and Mr. B. G. Churcher, Manager of the Research Department, and Sir Arthur P. M. Fleming, Director of

Research and Education, Metropolitan-Vickers Electrical Co. Ltd., for their permission to publish this paper.

REFERENCES

EITEL, W., and DIETZEL, A., 1946, *B.I.O.S. Report* No. 785, Appendix.
JACKSON, L. C., 1924, *Phil. Trans. Roy. Soc.* A, **224**, 1.
KENDALL, J. T., and SPRAGGON, W., 1947, *Report of XIth International Congress of Pure and Applied Chemistry.*
KENDALL, J. T., and YEO, D., 1948, *Nature, Lond.*, **161**, 476.
KRISHNAN, K. S., GUHA, B. C., and BANNERJEE, S., 1933, *Phil. Trans. Roy. Soc.* A, **231**, 235.
LONSDALE, K., 1937, *Rep. Prog. Phys.*, **4** (London : Physical Society), p. 368.
MIDDEL, V., 1946, reported by CURTIS, H. A., *Chem. and Met. Engng.*, March, 109.
NILAKANTAN, P., 1938, *Proc. Indian Acad. Sci.* A, **8**, 39.
RABI, I. I., 1927, *Phys. Rev.*, **20**, 174.
WEISS, P., 1905, *J. Phys. Radium*, **4**, 469, 847.
WILSON, E., 1920, *Proc. Roy. Soc.* A, **96**, 429.

The Scattering of Radio Waves

BY W. DIEMINGER
Institut für Ionosphärenforschung, Lindau über Northeim, Hannover, Germany

Communicated by W. R. Piggott; *MS. received 2nd December* 1949, *and in revised form 20th July* 1950

ABSTRACT. A review of the recent work on scattering draws attention to a number of discrepancies in the interpretation of the data. Most workers believe that the scattering is due to clouds in the E layer.

The experiments described in this paper indicate that there are several types of scattering phenomena. It is concluded that the most common source of scatter is the reflection of waves from irregularities on the surface of the earth. In this case the wave path is from the sender to F layer—earth—F layer and hence to receiver, i.e. is analogous to a 2×F reflection. The characteristics of the different types of scattering and their effects on the propagation of short waves are discussed.

§ 1. INTRODUCTION

THE most important mode of propagation of short radio waves depends on the regular reflecting properties of the ionized regions of the upper atmosphere, but there is now much evidence that scattered radiation is often an important factor in such propagation. This paper is concerned with phenomena due to the irregular structure of the ionosphere and the earth.

The amplitude and delay time of the echoes which are usually obtained during ionospheric sounding experiments vary comparatively slowly with time and the reflected pulse is generally a true replica of that emitted. However, in addition to these, echoes are sometimes observed which consist of a number of overlapping pulses, the amplitude and time delay of which vary rapidly. This produces on the screen of the cathode-ray tube a band of irregularly fluctuating pulses which are beating with a time constant between $\frac{1}{10}$ and 1 second. These echoes have been called scattered echoes, on the assumption that the abnormally rapid fluctuations are due to the presence of scattering centres.

The analogous phenomena in continuous-wave propagation (telegraphy and telephony) are the weak, fluctuating signals received inside the skip distance, i.e. in the area in which ray propagation is impossible. The field strength of these signals is one hundredth to one thousandth of the normal value and varies rapidly with time.

§ 2. PREVIOUS WORK

It appears that scattered echoes were first observed in 1926 by Mögel (Böhm 1929, Mögel 1932) monitoring the signals of the high power transmitter at Nauen. From the experiments with different types of aerials it was concluded that the radiation was reflected mainly vertically and that there were real reflecting layers at heights of 1,000 km. and more.

Independently, scattered echoes were observed by Eckersley in 1927 during experiments on high-power beam transmission. Eckersley concluded from very exhaustive experiments that scattered echoes are produced by clouds in the E layer on three different modes of reflection (Eckersley 1929, 1932, 1937 a, b, c): (*a*) 'E scattering'— direct scattering of the waves from the sender by E clouds (Figure 1(*a*)), (*b*) '1 × F scattering'— scatter of waves after reflection from the F layer by E clouds (Figure 1 (*b*)), and (*c*) 2 × F scattering by waves reflected in the F layer, scattered by E clouds and then reflected back to receiver by the F layer (Figure 1 (*c*)). It was believed that the results were inconsistent with scattering on the ground.

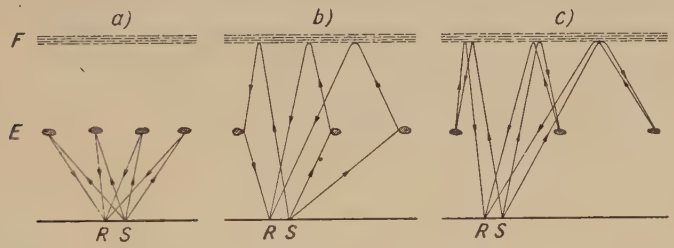


Figure 1. Different modes of scattering by E clouds : (*a*) E scatter, (*b*) 1 × F scatter, (*c*) 2 × F scatter.

Contrary to Eckersley's results, Edwards and Jansky (1941), using experimental data from long distance beamed transmissions, decided that some scattered echoes were produced by the inhomogeneous surface of the ground. They admit that other echoes are probably produced by irregularities in the ionosphere.

Experiments by Beckmann, Menzel and Vilbig (1939 a, b, 1940, 1941), carried out with a pulse sender of seven kilowatts, confirmed Eckersley's hypothesis as to the origin of 1 × F scatter. In addition the authors point out that the spread of echoes near the critical frequency observed at night may frequently be ascribed to the deviation by E clouds of the vertical trajectories to the F layer (Figure 1 (*b*)).

Netzer (1940) finds from routine (P',f) records, that scattering on frequencies near the critical frequency is correlated in 99% of the cases with a certain type of abnormal E ionization. He postulates a highly penetrating corpuscular radiation affecting both the E and F layer, i.e. genuine irregularities in both layers.

A quite different type of scattering was observed by Booker and Wells (1938) at Huancayo at the magnetic equator. At this station there are scattered echoes every night over a wide frequency band at ranges of 420 km. upward. The

authors assume that there is a region just above the maximum of the daytime F2 layer in which the electron density is inhomogeneous. That region is screened off by the F2 layer during daytime whereas during the night waves are scattered back from the irregularities.

This brief review of the published work on scatter phenomena clearly shows that different workers have studied different aspects of the subject and that there are considerable gaps in our knowledge both of the facts and of the reliability of the interpretations put forward in explanation.*

In particular it is important to discover the extent by which the results of different workers have been influenced by the techniques used, so that the reasons for the differences in interpretation can be established.

For these reasons an experimental investigation of scattering was planned at Lindau using as many techniques as could be provided with the limited facilities available, the experiments being planned, as far as possible, to provide solutions to the outstanding problems.

§ 3. EQUIPMENT

The data summarized in § 2 were obtained partly by using specially designed ionospheric sounding equipments with sender peak powers between 100 watts and a few kilowatts and partly by modulating high-power (50–100 kw.) communication transmitters with pulses for particular experiments.

The results obtained by the former are seriously incomplete since the power was usually too small to give scattered echoes of adequate amplitude regularly. The latter could only be used for comparatively short periods between commercial transmissions and were not available for continuous experiments.

For our experiments at Lindau special equipments were constructed which could be used for unlimited periods. The technical data of these equipments are summarized below.

A. *High power constant frequency sender.* Self oscillator sender using two high emission triodes with anode modulation, adjustable to any fixed frequency between 3 and 20 Mc/s. Pulse modulation: 50 pulses per second with pulse duration 10^{-4} second. Aerial: Horizontal six wire cage dipole length $2 \times \lambda/4$ at $\lambda/4$ height above ground. The aerial was adjusted in length and height for each frequency used. Peak power in aerial 50–100 kilowatts.

B. *Automatic P',f sender.* Push-pull master oscillator–power amplifier sender using 50 watt pentodes in the oscillator and 800 watt pentodes in the output stage, anode modulated throughout. Frequency range 1–16 Mc/s., the sender being automatically tuned to a stable receiver. Pulse modulation: 50 pulses per second with pulse duration $0{\cdot}8 \times 10^{-4}$ seconds. Aerial: V-shaped cage dipole diameter 1·7 m., length 2×50 metres. Peak power in aerial 10–20 kilowatts.

C. *Receivers.* Superheterodyne receivers of conventional pulse type with bandwidth 15–20 kc/s.; sensitivity adequate to give at least 10 volts output from the receiver noise on all frequencies.

D. *Presentation units.* Conventional cathode-ray tube recorders with linear time-base and automatic frequency, time and range calibrations. 'A' scope

* *Note added in proof.* The problem of scattering has recently been investigated by A. H. Benner (*Proc. Inst. Radio Engrs.*, 1949, **37**, 44), A. M. Peterson (*J. Geophys. Res.*, 1949, **54**, 284) and R. Silberstein (*J. Geophys. Res.*, 1949, **54**, 288). The results agree with the conclusions drawn in this paper.

display for visual monitoring and snap photos of echo pattern and intensity modulation for continuous and automatic P', f recorders. All records are made on 35 mm. film.

§ 4. TYPES OF SCATTERING

The experiments carried out with these equipments prove that there are several types of scattering, differing in appearance, frequency and apparently in origin. From the ranges of the scattered echoes it is possible to discriminate between (i) E scatter, (ii) F scatter, (iii) G scatter, (iv) $2 \times$ F scatter.

It is immediately evident from even a preliminary examination of the records that the $2 \times$ F scatter is the most frequent and, under undisturbed conditions, the most important one. Therefore the investigations to be described are mainly devoted to this phenomenon. The other types observed in the course of the experiments are reviewed briefly for the sake of completeness.

(i) E *Scatter*

The E scatter exists in two forms. The first consists of echoes having a duration of the order of one second with a typical rapid increase in echo amplitude followed by a slower decrease. These appear at ranges between 90 and 200 kilometres and show the same characteristics as the meteor echoes described by Appleton and Piddington, Eckersley and many others. It appears to be certain that these echoes are caused by meteors producing short-lived ionized trails in the E region of the atmosphere. The longer ranges appear to be due to distant trails in E region reflecting obliquely. Further data and references are given in a paper by the author (1947). Extremely large numbers of these echoes are recorded at about 100 km. height using the high power sender on frequencies between 3 and 4 Mc/s. both by day and night and this phenomenon deserves further investigation. It is well known that the frequency of observation of meteors increases rapidly with decreasing size and it can be shown that there must be a nearly continuous stream of very small dust particles impinging on the upper atmosphere which should produce a detectable amount of ionization. It appears probable that the echoes observed on these low frequencies may be due to this cause.

The second kind of E scattering shows a broadened echo pattern, the upper limit of which may reach to 200 km. (Figure 2, Plate I*). The phenomenon tends to occur most frequently at sunrise and sunset and the echo pattern is similar to that observed at high latitude stations during auroral disturbances in E region.

A few short references to this type of phenomena have appeared (Appleton, Naismith and Ingram 1936, Eyfrig *et al.* 1938, Beckmann, Menzel and Vilbig 1939 b, Netzer 1940) but no exhaustive investigation into its occurrence and origin has been published.

(ii) $1 \times$ F *Scatter*

The $1 \times$ F scatter phenomenon is shown in Figure 3 (Plate I*). The splitting of the ray into two components is marked by a great number of traces S penetrating near the critical frequency. Sometimes the phenomenon extends over a band of 2 Mc/s. This type of reflection is less marked, though still observable, in the multiple reflections. $1 \times$ F scatter is more common at night than in the day and is particularly frequent during winter nights.

* For Plates see end of issue.

As mentioned in §2 above, Beckmann, Menzel and Vilbig have attempted an explanation of this type of scattering in terms of an increased angle of incidence at the layer due to laterally displaced scatter centres in E region. Rays striking the F layer at the steeper angles will penetrate the layer at lower frequencies than rays incident more obliquely. The results given by Netzer may be explained by the same mechanism.

This explanation appears to have been adopted by workers at the Central Radio Propagation Laboratory (U.S.A.) who state that when spread echoes are present in the F layer the 'inner edge' (lowest penetration frequency) should be used for measuring the F2 critical frequency. It must be pointed out that this explanation does not explain the diurnal and seasonal frequency of occurrence of $1 \times F$ scatter satisfactorily nor the well marked $1 \times F$ scatter which is present in some types of ionospheric storm.

It is possible to prove conclusively that deviation by E clouds is not the only source of $1 \times F$ scatter. If there is a deviation in E region capable of causing a spread of F-layer critical frequency, then on the lower frequencies trajectories must be present giving a similar spread of height (owing to the long deviated path involved). Hence on this mechanism the spreads in height and critical frequency should closely correspond. Examination of the experimental data shows that the theoretical corresponding spreads do, in fact, occur on a large number of occasions but that there are many occasions where there is a well marked $1 \times F$ scatter near the critical frequency and no spread of height on the low frequencies. We must conclude that, on these occasions, the $1 \times F$ scatter is due to genuine irregularities in the F region. These cases include, of course, many occasions during ionospheric storms when $1 \times F$ scatter is seen very distinctly.

Hence it should be remembered that spreading both in critical frequency *and* in height at lower frequencies may also be explained in terms of turbulence in the F region.

There is no generally accepted hypothesis explaining the seasonal variation of $1 \times F$ scatter at night. Our observations suggest that the winter maximum is due to true F region turbulence which produces its maximum effect when the layer relaxation time is greatest, that is in winter nights. Further data are needed to establish the truth of this suggestion.

(iii) G *Scatter*

In agreement with the usual nomenclature used in ionospheric work, scattered echoes which are separated distinctly from the F scatter and which have a minimum apparent height greater than that of the F layer will be called G scatter.* G scatter appears on a (P', f) record as a cluster of echoes at 600–1,500 km. minimum height. The lower limit in height is defined rather well but there is no sharp boundary towards the top of the record. The minimum height varies between 600 and 1,500 km. for different times of observation. The echoes are observed on frequencies above the F2 critical frequency, covering a frequency-range up to several megacycles per second above the critical frequency. They weaken at high frequencies without marked change in the apparent height but there is a distinct increase in height at the lower frequencies. For some 50% of the time when G scatter is present there is an overlap between the F2 echoes rising towards the

* *Note added in proof.* The same phenomenon has been observed independently by R. Rivault (*Proc. Phys. Soc.* B, 1950, **63**, 126).

critical frequency and the G scatter rising towards lower frequencies. A typical example of this type is shown in Figure 4 (Plate I). Under favourable conditions G scatter may be observed rising to ranges up to 2,000 km. apparent height towards lower frequencies. G scatter mainly occurs after midnight at Lindau and the minimum apparent height tends to decrease through the night. On some, but not on all, occasions G scatter is coincident with an auroral type ionospheric disturbance. There is little doubt that G scatter is identical with the 'Nordlichtschicht' observed many years ago (Dieminger and Plendl 1938, Beckmann, Menzel and Vilbig 1938 a, b) on fixed-frequency records. The overlap of the F ray and G scatter patterns proves conclusively that the G scatter involves lateral deviation and hence it is not possible to derive, conclusively, the ionospheric structure from observations made at a single station.

The characteristics of G scatter are very similar to those of the equatorial type of spread echoes (Booker and Wells 1938) as observed during sunrise. Hence it may be concluded that the mechanism of G scatter is identical with that proposed by Booker and Wells for the equatorial type. But there are two differences: (i) At Huancayo spread echoes are present every night, whereas in Central Europe G scatter is a rather rare phenomenon. For example, in February 1948 G scatter has been observed in 58 of 1,390 observations, i.e. about 4% of the total observation time. (ii) The overlap of the F ray and of the spread echoes at Huancayo during sunrise is ascribed to the gradient in ionospheric conditions in an east–west direction due to the difference in the elevation of the sun. For G scatter it appears to be a gradient in a north–south direction owing to the different distance from the auroral belt.

During auroral disturbances the electron density of the F2 layer decreases towards the auroral belt and at the same time spread echoes are observed, the height of which decreases in the same direction. These echoes appear to be due to electron clouds. A (P',f) record taken at a point in the south will be undisturbed. At a point near the auroral belt the lower frequencies will be reflected normally by the F layer whereas the higher frequencies will penetrate the layer and be scattered by the irregularities above and in it. The height of these scattered echoes will be nearly independent of frequency for all frequencies well above the critical frequency. Near the critical frequency the delay of the scattered echoes will increase due to retardation in the F layer. In addition there will be obliquely scattered echoes on frequencies below the critical frequency which have penetrated the layer in a northerly direction where the electron density is lower than just above the point of observation. The tilt of these rays, and therefore the time delay of the echoes, will increase with decreasing frequency. Finally at a point within the auroral belt, the well-known auroral type of disturbance will be observed.

The variation of minimum apparent height during the night is interpreted as being caused mainly by the movement of the disturbance towards the south. A final proof of this hypothesis may be obtained by a comparison of the records from several stations suitably spaced along a meridian.

(iv) $2 \times$ F *Scatter*

Using a high power pulse sender, the $1 \times$ F reflection under normal conditions appears on the oscilloscope screen as a true replica of the ground pulse, whereas the $2 \times$ F and higher multiples are followed by a number of echoes the amplitudes of which decrease exponentially with time delay. On a (P',t) record (Figure 5,

Plate I) there is no widening of the $1 \times F$ trace whereas the $2 \times F$ extends nearly to the $3 \times F$ reflection. Similar effects are visible, with decreasing amplitude, on the higher multiples. The dependence of the appearance of the scatter on the power used is shown very distinctly in Figure 5 (Plate I) in which the receiver sensitivity is varied between several different values. The amplification is greatest during the periods *c*, medium during *b* and lowest during *a*. It is very significant that there are no scattered echoes below the $2 \times F$ level at any of these levels of sensitivity.

Some evidence bearing on the source of the scatter is given by its behaviour near the F layer critical frequency. The (P', t) record (Figure 6, Plate II) shows the well-known splitting into the ordinary and extraordinary components and the penetration of both for the $1 \times F$ and $2 \times F$ reflections. The scattered echoes do not follow the upward trend of the ray and at the time of penetration the scatter is distinctly lower than the $2 \times F$ ray reflection. The scattered echoes continue to exist some time after the ray penetration but disappear eventually as the electron density decreases.

The following are the main features of the phenomenon as shown by large numbers of records: (i) there are no scatter echoes below the $2 \times F$ ray reflection on frequencies sufficiently below the critical frequency; (ii) the amplitude of the scattered echoes is almost independent of frequency; (iii) the $2 \times F$ scatter exists regularly and can always be observed if sufficient power is used. It is not a sporadic disturbance.

§ 5. THE SOURCE OF $2 \times F$ SCATTERING

It is obvious that if senders of infinite power were available scatter would be obtained from every irregularity in the ionosphere and on the ground and every mechanism which is physically possible would contribute to the scatter pattern.

We shall therefore attempt to explain our observations in terms of one simple mechanism and, if such an explanation proves to be possible, we shall interpret the result as indicating that this mechanism is more efficient and gives a greater amplitude of scatter than any alternative mechanism.

Two explanations of $2 \times F$ scatter are possible: (i) that it is due to reflection in the F layer, the scattering being from E clouds, (ii) that it is due to reflection in the F layer, the scattering occurring at the surface of the earth.

(i) *Scattering by* E *Clouds*

First we shall attempt to explain our results in terms of Eckersley's mechanism involving scattering from E region clouds.

If the scattering takes place from E region clouds (Figure 1) then, as the angle of incidence is decreased to zero, the limiting value of the minimum apparent height of reflection of the scatter will be $2h'_F - h'_E$. In practice we find that, apart from the limited frequency range near the critical frequency, the markedly greater height of $2h'_F$ is observed in every case. Eckersley avoids this difficulty by assuming that the amplitude of the waves scattered back at nearly vertical incidence is below the noise level of the receiver. If this assumption is true the lower boundary of the scattered echoes should vary in height when the sensitivity of the receiver is altered. Actually it is observed (Figure 5) that only the upper boundary depends on the sensitivity used. When the amplification is reduced the long delay scatter vanishes first and the spread of the $2 \times F$ echo decreases continuously until finally only the ray reflection is seen on the oscilloscope.

Similar evidence is given by the records showing M reflections (sender–F–E–F–receiver). These echoes are very frequently seen simultaneously with, or sometimes a short time before or after, strong abnormal E reflections. It is generally accepted that these reflections are due to irregularities in the E region and the path involved is equivalent to that postulated by Eckersley to explain $2 \times F$ scatter. If the $2 \times F$ scatter is, in fact, produced by this mechanism it should be strongest when M reflections are present and the lower boundary of the scatter should coincide with the height of the M reflections. It is found experimentally that there is occasionally some weak scatter associated with the M reflection but that, even in these cases, the main scatter is associated with the $2 \times F$ reflection. There are very few cases in which the amplitude of the M scatter is greater than that of the $2 \times F$ scatter and even for these cases the $2 \times F$ scatter is markedly dominant on frequencies above the critical frequency. As far as can be deduced from the small amount of data available—records showing any M scatter being extremely rare—M scatter is associated with spread E echoes and the cause appears to be a special and unusual stratification of the E layer. It is probable that the records discussed by Beckmann, Menzel and Vilbig (1939 a) were taken during these conditions.

We may conclude that the E cloud hypothesis of Eckersley is in disagreement with the experimental evidence.

(ii) *Scatter from the Ground*

Now we shall consider the phenomenon on the assumption that the waves are scattered back from irregularities of the ground.

Assuming, as a first approximation, that the height of reflection does not vary with the angle of incidence, then the lower limit of the delay of the scattered echoes will increase with the obliquity of the rays and its minimum value will correspond to that of a $2 \times F$ ray reflection. This is in good agreement with the experimental observations.

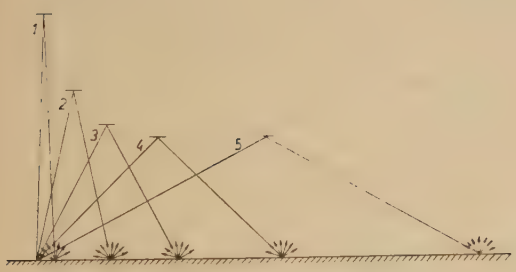


Figure 7. Paths of rays of different angle of incidence near the critical frequency.

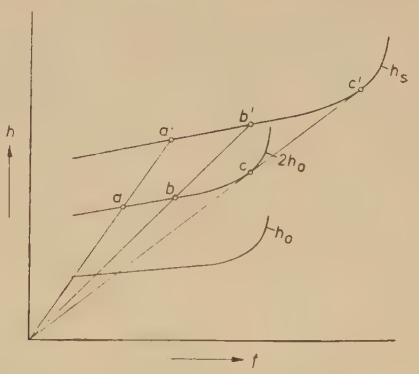


Figure 8. Minimum heights of scattered echoes derived from the vertical (P', f) curve.

In practice, near a critical frequency, the height of reflection will vary with frequency and will decrease with increase in the angle of incidence. The main features of the phenomenon are shown qualitatively in Figure 8. It is evident that paths of medium obliquity, for example 2, 3 in Figure 7, will take less time than those at vertical incidence, that is near the critical frequency the path of the

scattered echoes will be shorter than that of the $2 \times$ F ray reflection. For very oblique rays the increase in range dominates the decrease due to lowered height of reflection, and the apparent delay increases.

The minimum apparent height of scatter h_s may be computed by a simple procedure from the shape of a (P',f) curve. If $(h_s)_\theta$ is the apparent height of scatter at angle of incidence θ, f_θ is the highest frequency which would just be reflected at this angle, and h_0 and f_0 are corresponding heights and frequencies at vertical incidence, then it can readily be shown that $(h_s)_\theta = 2hf_\theta/f_0$.

Thus from a given vertical incidence (P',f) curve, the curve $(h_s)_\theta$ for f_θ/f_0 a constant may be obtained by expanding the heights and frequency scales by the factor $f_\theta/f_0 (=\sec\theta)$: in Figure 8 point *a* becomes *a'*, *b* becomes *b'*, etc. The new curve gives the height of the scattered echo as a function of frequency for a given angle of incidence. This is not identical with the (P',f) curve for oblique incidence which gives the corresponding relation for a fixed distance and not for a fixed angle. Change in the angle of incidence produces a family of curves representing all possible paths for the scattered echoes (Figure 9). The lower limiting curve to this family, which represents the curve of minimum apparent range of the scatter as a function of frequency, is given partly by the curve $h = 2h_0(f_0)$ and partly by the tangent from the origin to this curve. Clearly, near the critical frequency the minimum heights of the scatter must be less than the height of the ray reflection. Rigorously, of course, this construction can only be used when the ionosphere is homogeneous over an adequate area and care must be used when horizontal ionization gradients are possible.

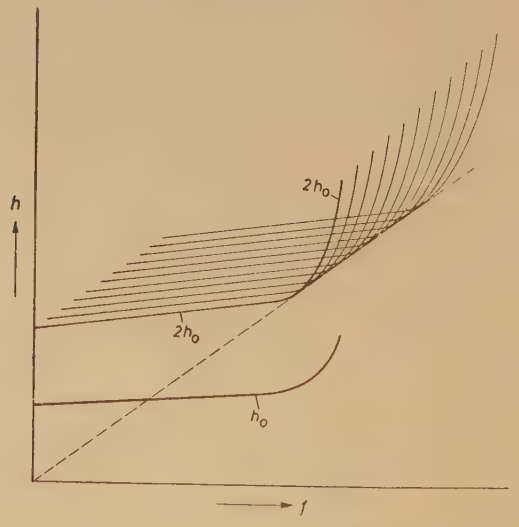


Figure 9. As Figure 8 but with family of curves.

If the scatter is due to E clouds the same procedure gives another limiting curve made up of the M reflection and the tangent to the M curve, $h = 2\,h_{0F} - h_{0E}$, and the path of the echoes would be markedly shorter.

Before comparing the computed and observed values of minimum apparent height we have to consider the effect of the earth's magnetic field in the ionosphere. The theory given above will apply to the ordinary component 'o' whereas the highest frequency reflected will be due to the extraordinary component 'x'. At

vertical incidence at these latitudes the separation between the 'o' and 'x' critical frequencies is approximately 0·7 Mc/s. At oblique incidence the corresponding relation is a complex function of the angle between the direction of propagation and the magnetic field which varies with the geographical position of the stations and with the distance. However, as a rough approximation for the shorter distances we can displace the pattern 0·7 Mc/s. towards higher frequencies. In Figure 10 the comparison between computed and observed scatter ranges is carried out for a typical record. The curves show both the predicted 2 × F scatter (dashed) and the M scatter (chain line). The observed values shown by crosses fit much better with the former and there is no doubt that this favours the assumption that the scatter comes from the ground.

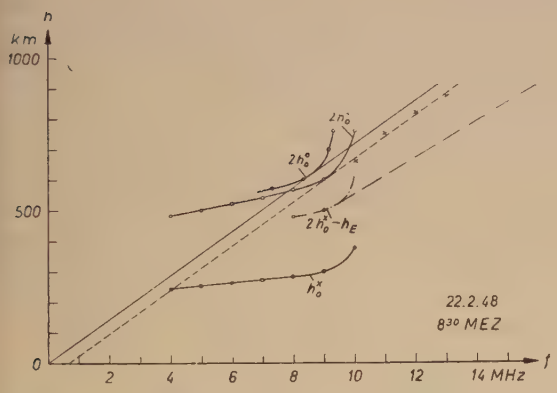


Figure 10. Computed (– – –) and observed (× ×) minimum heights of scattered echoes.

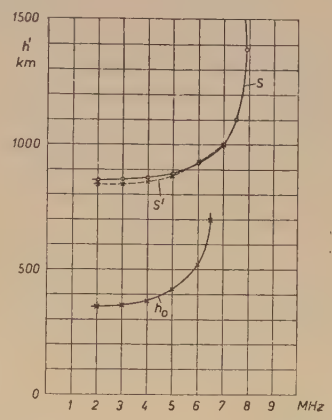


Figure 13. Computed (P', f) curve of a re-radiator at 500 km. distance compared with the observed curve of maximum intensity of scattered echoes.

If the 2 × F scatter is due to irregularities of the ground similar phenomena should be present near the 2 × E reflection. In fact we must expect the 2 × E echo trace to be broadened upwards and to persist to a higher frequency than that of the 1 × E trace. As far as is known, this effect has never been observed in the case of reflections from the normal E layer, probably because the normal E layer is only present during the daytime when absorption is great. This would cause the amplitude of the scattered signal to be small compared with the noise level at the receiver. Thus during the daytime the 2 × E reflection is usually lost by absorption below the 1 × E critical frequency. However, during the night, when absorption is low, the expected effect is frequently observed (Figure 11, Plate II) and we find: (i) that there are scattered echoes above the 2 × E ray reflection; (ii) that the 2 × E reflection extends to higher frequencies than the 1 × E trace; (iii) that the height of these 2 × E reflections slowly increases towards the higher frequencies, this increase corresponding to the longer paths of waves reflected obliquely.

§ 6. THE SOURCES OF THE SCATTER FROM THE GROUND

The existence of ground scattering on medium wavelengths has been established by Gerber and Werthmüller (1940) who measured the field near a broadcasting station and showed the presence of interference between the direct radiation and that scattered from a woody hill.

The scattering may be due to the following mechanisms: (i) Rayleigh scattering produced by irregularities which are small compared with the wavelength used, e.g. irregularities of the ground, bushes, waves on the sea, etc. This corresponds to a surface which reflects diffusely, and the intensity at very oblique incidence will decrease very rapidly.

(ii) Re-radiation from obstacles whose size is of the order of one wavelength, e.g. buildings, trees, poles etc. These may be regarded as reflectors excited by the incident radiation. Currents have been observed in trees in the fields of broadcasting stations (Gerber and Werthmüller 1940). The polar diagram of the re-radiation depends on the shape of the reflector and on the ratio of its length to the wavelength.

(iii) Reflection by tilted surfaces the dimensions of which are large compared with a wavelength, e.g. hills and mountains. Once again there is a statistically large reduction in the back re-radiation at low angles of elevation owing to the relative uncommonness of steep slopes. In all cases the dielectric properties of the ground cause a large reduction of the re-radiation at low angles of elevation.

For a constant frequency below the critical frequency, in which case scatter will be receivable at all angles of incidence, we may note that the three modes of scattering should give different variations of amplitude with range.

Consider observations made during nights when ionospheric absorption is negligible. The scatter amplitude will vary inversely with the apparent range and will vary with the angle of incidence θ at the sender, receiver and scatter point. After allowing for the receiver and sender polar diagrams, which may be readily calculated, the remaining factor is controlled by the scatter mechanism.

In the case of Rayleigh scattering from small centres on the ground the polar diagram of the scatter centres must lie between the functions $\sin\theta$, $\cos\theta$ for vertically polarized centres and $\cos\theta$ for horizontally polarized centres, and hence the residual variation of amplitude with range z is controlled by a function bounded by $\sin^2\theta\cos^2\theta$ and $\cos^2\theta$. Now $\cos\theta \simeq 2h_0/z$. This gives a maximum variation of amplitude with range due to this factor which is inversely proportional to the square of the range. Since the number of centres at a given range is likely to increase with range a somewhat slower variation would be expected experimentally.

For scatter due to scattering centres whose size is comparable with the wavelength used a similar conclusion is found, the variation of scatter with distance being of the same type but rather slower than for the Rayleigh scattering owing to the influence of the longer obstacles in concentrating the scattered energy at the larger angles of incidence.

In the case of scatter from areas large compared with a wavelength the variation of scatter amplitude with distance is primarily controlled by the number and area of surfaces with a given tilt. An examination of large scale maps shows that areas with very small angles of tilt are overwhelmingly predominant and we should therefore expect that, for this case, the amplitude would decrease extremely rapidly with range. Experimentally, it is found that the amplitude of scatter does, in fact, decrease much more rapidly than would be expected from the first two mechanisms. For the variation in range over which scatter is observed on these frequencies the polar diagrams of the sender and receiver can be regarded as circular, that is, the equipment sensitivity is independent of θ in this range.

Above the critical frequency this approach is no longer justified since there are serious focusing phenomena at the edge of the skip distance which modify the variation of amplitude with range at constant frequency.

Regardless of the actual mechanism of scattering energy will be scattered from a great many points surrounding the point of observation. The resultant pattern will be made up from superposed echoes of constant delay but different azimuths, i.e. echoes reflected from points on a circle centred on the receiver, and from echoes having different time delays because reflected from different distances. Continuous small changes in the ionosphere will make the phase delays of all these echoes fluctuate and will therefore prevent a stationary interference pattern from being formed. Thus the result will be a broad band of echoes rapidly fluctuating in amplitude, the position of the instantaneous maxima and minima being controlled by the random phases and amplitudes of the individual scattered echoes.

In general, the amplitude of the scatter echoes is controlled by three functions which vary with the range: (i) the field strength decreases as the inverse of the apparent range of the signal due to spatial attenuation; (ii) the echoes at longer ranges are scattered with greater angles of incidence and the back scattering function is smaller; (iii) the absorption in the D layer of the ionosphere increases with the obliquity of the rays in this layer and is therefore greater at the longer ranges.

Theoretically it should be possible to distinguish between the different scattering mechanisms discussed above but insufficient accurate measurements of the variation with frequency of the amplitude of the scattered echoes are available at present to justify elaborate calculations for this purpose. Nevertheless, as has been shown above and by Eckersley and others, there is sufficient evidence to exclude the possibility that the scatter is predominantly of the Rayleigh type and some features of the scatter strongly suggest that large irregularities of the ground give rise to most of the signals.

§7. INFLUENCE OF THE SHAPE OF THE GROUND

If the macroscopic features of the ground are dominant sources of scatter there should be traces in the scatter associated with large geographical features such as mountains. Such features are actually found in the records observed at Lindau/Harz. There is a distinct maximum in the amplitude of the scatter for echoes at ranges corresponding to ground distances of 400-600 km.; this is the distance from Lindau to the Northern slopes of the Alps.

On a (P',f) record these echoes from an almost continuous trace lying between the $2\times$F and $3\times$F ray reflections over a wide range of frequency (Figure 12, Plate II). The height of these echoes increases with frequency much more slowly than the height of the $2\times$F ray reflection. This would be expected for a reflection from a distant source.

Theoretically, the height–frequency curve should be similar to an oblique incidence (P',f) curve since the scattering surface is equivalent to a re-radiating sender. Thus the theoretical delay curve may be easily computed from the vertical incidence (P',f) curve using conventional methods for the double path. In Figure 13 a typical example is computed for the mean Alps–Lindau distance of 500 km. The solid curve S is calculated from the observed vertical incidence (P',f) curve and the dashed curve S′ is drawn using the observed time delays of the strong Alp echoes from the frequencies for which it is distinct from the general scatter. The difference between the curves is less than the inaccuracies of measurement.

Though common, these strong Alp echoes are not distinct in all records. A probable explanation of this may be given in terms of the small ripples and

irregularities which, it is well known, occur in the ionosphere. Seen from the ground the ionosphere near the reflection area appears alternately concave and convex and, depending on the geometry of the reflection, there will be a focusing or defocusing of neighbouring rays at a given distance. For scattered radiation the resultant changes in amplitude are very marked since the effect is roughly doubled by the double reflection at the ionosphere. According to the position of the ripples different areas of the earth's surface will contribute dominantly to the scatter, and the shape of the scatter pattern will be changed slowly by the movement of the irregularities in the ionosphere. With omni-directional aerials these effects are partly veiled by the scatter from other directions. It is intended to extend these observations using highly directional aerials.

§8. THE INTENSITY OF SCATTERING AS A FUNCTION OF FREQUENCY

The amplitude of scattered echoes observed at any given frequency depends on the range of the scattering sources, on the actual mechanism of scattering, on ionospheric absorption and on the polar diagrams of the sender and receiver aerials. Experimentally it is found that the amplitude of the scatter signals decreases towards lower frequencies, due primarily to ionospheric absorption, whilst the upper frequency limit is determined by sender power and receiver sensitivity and, in the case of very high sender powers, on the aerial polar diagrams. Daytime observations show that at the higher frequencies the amplitude is almost independent of frequency up to a value some 30% above the normal incidence critical value and then decreases rapidly.

There is a noticeable difference between the amplitudes of the scattered signals on the two sides of the $2 \times F$ ray reflection, those below or at higher frequencies than the ray being little stronger. This is due to the existence of two paths of propagation in the latter case—the normal and the Pedersen high angle path —whereas only one path is possible for reflections below the critical frequency. This increases the mean amplitude by $\sqrt{2}$, since the phases are random, and can give greater increases where the rays coincide at the skip distance.

§9. TEMPORAL VARIATIONS IN SCATTERING

The hypothesis that the scatter is due to reflections at the ground does not involve any special ionospheric condition apart from the limitation that absorption must not be excessive. Hence one would expect that $2 \times F$ scatter should be present regularly and, if it is not observed, a reason for its absence is essential. This test has been applied to the data available and a typical example of the results may be quoted here. In February 1948 out of a total of 1,053 observations the number of observations with no $2 \times F$ scatter was 16.

Of these 16 cases 6 were coincident with a period of reduced power at the sender such that the scattered energy would be expected to be less than the noise input to the receiver. The remaining 10 cases all occurred at noon on days of high absorption and are fully explained by the absorption present. Further observations during 1949 showed that $2 \times F$ scatter can always be obtained provided the sender power is adequate.

The $2 \times F$ scatter is least marked during periods of strong $1 \times F$ scatter. This may be due either to multiple scattering effects in the layer dissipating the energy or to the effects of the layer tilts which are commonly associated with $1 \times F$ scatter. When $1 \times F$ scatter is due to E clouds the energy in the $2 \times F$ scatter trajectory will

be scattered five times—once at the ground and four times in the E layer. It is found experimentally that little or no $2 \times F$ scatter is received when there are marked irregularities of this type in the E region.

§ 10. SCATTERING AND RADIO COMMUNICATION

(i) *The Influence of Scattering on Reception and Direction Finding*

The practical importance of the scattering in short wave propagation has been repeatedly stressed by Mögel, Eckersley, Beckmann, Menzel and Vilbig and by Grosskopf and Vogt (1940). It is shown that, above the maximum usable frequency the mode of propagation changes from ray reflection to scattered radiation. The field strength of the latter is 40-60 db. below the field strength of the ray reflection and the signal is subject to rapid and violent fading. Nevertheless the field strength is sufficient for communication with high transmitter powers of the order of 100 kw. The transition from ray to scattered signal shows characteristic differences depending on whether the ray reflection is from the abnormal E or the F layer. For the former there is a gradual transition due to variable and ill defined abnormal E critical frequency and for the latter there is striking interference fading between the ordinary, extraordinary and Pedersen rays followed by an extremely rapid decrease in signal strength when the ray transmission fails.

The effects of scatter phenomena on radio communication depend fundamentally on whether the receiving station is within or beyond the skip distance from the sending station. So long as the skip is shorter than the distance between sender and receiver (Figure 14(*a*)) ray propagation is possible and the field due to the scatter will be 100 to 1,000 times less than the field due to the ray and may be neglected. In the case of direction finders using the minimum signal method the scatter is, however, very important since the observation is made when the direction finder has been adjusted to eliminate most of the ray reflection. Thus the depth of the minimum, and the sharpness and position of the bearing will be affected by the strength and direction of the scatter. Scatter thus imposes an absolute limitation to the accuracy obtainable even with perfect equipment on a perfect site. For example, if the ratio of the amplitudes of the scatter to the ray is 0·01, the former will exceed the latter over a range of 35′ about the minimum, i.e. the maximum bearing sharpness will be about 1°.

When the skip distance exceeds the sender–receiver distance, $2 \times F$ scatter can only reach the receiver from points outside the scatter curve defined by the outer envelope of the skip distance curves around the sender and receiver, i.e. from the hatched area in Figure 14 (*b*) and (*c*). Thus when the skip distance increases through the sender–receiver distance there is a rapid increase in the delay of the signal which must now travel from the sender to the scattering point P and hence to the receiver instead of by the direct path sender to receiver.

Although scattering may occur from any point outside the scatter curve there will be a predominance of radiation from the areas near the intersections of the two skip-distance curves. This is caused partly by geometrical factors, these areas giving the shortest paths with the smallest angles of incidence, and partly by the focusing phenomenon which occurs along the skip curves and which will be greatest where two skip curves intersect.

The effect may be demonstrated very clearly by changing the sender frequency continuously and observing over a fixed distance (Figure 15). At low frequencies

the 1 × F and 2 × F ray reflections are present. At the frequency for which the skip distance is half that from the sender to receiver the 2 × F ray ceases (f_g''), leaving only 2 × F scatter and 1 × F ray. When the skip distance is equal to distance between sender and receiver the 1 × F ray reflection also ceases and only the 2 × F scatter remains (f_g'). Equivalent results are obtained using a fixed frequency during a time of changing ionization density, for instance, in the morning or evening.

Several experiments of this type have been made over the 700 kilometre east–west path between Lindau and Slough, England. These show the effect very clearly. In the morning the 2 × F scatter first appears at great ranges and its delay gradually decreases. The 1 × F ray reflection with its Pedersen ray then appears suddenly at a much shorter range and the 2 × F scatter range continues to decrease until the 2 × F reflection appears in it. Details of the corresponding direction-finding experiments are being published elsewhere by members of the Radio Research Station, Slough. It may be stated here that these results are in excellent agreement with the proposed mechanism shown in Figure 14, the 2 × F scatter before the 2 × F ray shows fluctuating bearings, the mean direction of which is very different from the great circle direction of Lindau from Slough. The

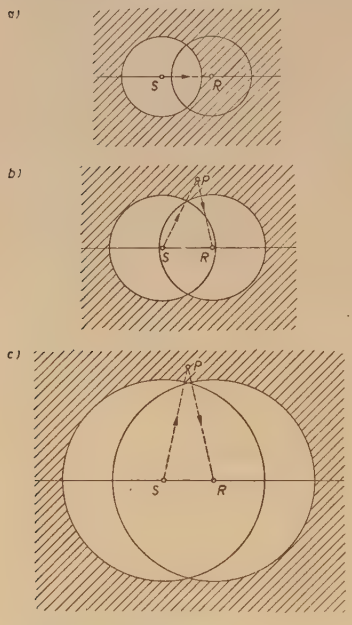


Figure 14. Paths of propagation for different ranges of the skip distance. S, sender; R, receiver; P, point of re-radiation.

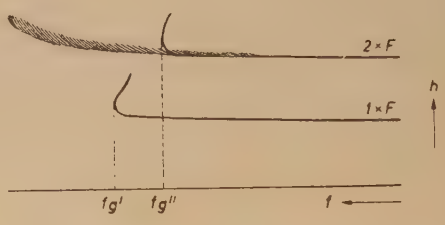


Figure15. (h', f) curves for 1 × F and 2 × F with f_g' and f_g''.

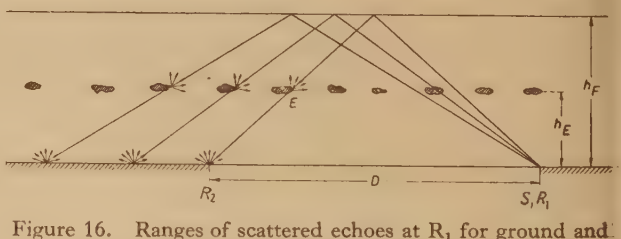


Figure 16. Ranges of scattered echoes at R_1 for ground and E reflection when the skip distance passes the receiver R_2.

bearing of the 1 × F ray reflection is correct. In the Lindau–Slough experiments the change from scatter to ray reflection was very clearly marked by the bearing changes and it was possible to fix the time when the ray appeared to within a fraction of one minute. The main scatter normally came from the south. This is not surprising since the possible scattering area to the north is sea whereas to the south the scattering area lies in the French hills. Furthermore the variation of skip distance with azimuth was such as to favour scatter signals from the southerly direction.

It may be noted that a further experimental test of whether the scatter is from the ground or from E region may be obtained using the direction-finding experiments. From the jump in bearing when the ray appears it is possible to determine the time when the skip boundary passes through the receiving station (R_2 in Figure 16) with great accuracy. If the scatter comes from the ground the scatter echoes with the shortest delay observed on the receiver R_1 situated near the sender, must come from R_2, whereas, if the scatter is produced by E clouds it will come from point E.

In the first case the echo range will be $S=2(h_F^2+\frac{1}{4}d^2)^{1/2}$ whereas in the second case $S=(2-h_E\,h_F)(h_F^2+\frac{1}{4}d^2)^{1/2}$. The difference between these values (about 20%) is large enough to provide a clear distinction between the hypotheses. In order to be convincing it is necessary to be certain that the minimum skip distance is in the direction of the receiver. For this reason it is necessary to carry out the experiments at times when ionospheric conditions in the required direction are favourable and it is also desirable to use directional aerials to limit the possibility of obtaining misleading results when there are abnormal conditions in the ionosphere.

(ii) *Limits of the Reception of Scattered Radiation*

When the frequency is increased the skip distance increases also and the angle of incidence of the waves increases. Eventually a frequency is reached above which no reflection is possible. This is known as the limiting maximum usable frequency.

No $2\times F$ scatter is observable above this frequency. In practice aerial sensitivity decreases rapidly at large angles of incidence and this limits the maximum angle of incidence and hence the maximum frequency at which scatter can be observed.

In past experiments the longest ranges at which $2\times F$ scatter was observed were between 4,000 and 4,500 km. From the geometry of the ionosphere the corresponding maximum usable frequency factor must be less than about 3, and the area of reflection will be at a range of about 2,000 km. Thus $2\times F$ scatter echoes from a given direction on a given frequency will vanish when the critical frequency in an area 2,000 km. distant in the required direction falls below about one third of the working frequency. Also if the critical frequencies at all points in the area of 2,000 km. radius round the sender or receiver are less than about one-third of the working frequency no $2\times F$ scatter will be observed.

This hypothesis has been tested qualitatively during winter nights and has been found to account for the main variations of the maximum frequency of observation of $2\times F$ scatter. It is hoped to make quantitative tests when ionospheric data are available from more stations in Europe.

This effect has an important control on the amount and type of scatter observed at different times on high-power high-frequency commercial transmissions and has undoubtedly contributed to the contradictions concerning scattered signals in the literature.

(iii) *Scatter as a Mode of Propagation*

Since $2\times F$ scatter is always present in medium and high power ionospheric sounding experiments it should be a permanent feature of high power radio communication. This, in fact, is found to be true. For example a study of a large number of field strength records taken between 1940 and 1945 using short-wave transmissions in the band 6 to 10 Mc/s. over ranges of about 1,000 km.

showed that the field strength never fell to zero when the skip zone reached the receiver but fell to about one hundredth of the normal and then decreased slowly with time. However, there are large differences in the amplitude and steadiness of the residual field, since, under disturbed conditions, the normal trend of $2 \times F$ scatter is obscured by other modes of scattering.

The possibility of using scattered signals for communications purposes depends on the field strength and its steadiness, that is, the quality of transmission required at the receiver. Since the effect of $2 \times F$ scatter is regular and controlled by the recognized properties of the ionosphere it is possible to predict the workable frequencies.

ACKNOWLEDGMENTS

This work was carried out in the Institut für Ionosphärenforschung in the Max Planck Gesellschaft. The author wishes to express his thanks to K. H. Geisweid for his help in building the equipment and carrying out the experiments, to R. Eberhard and H. G. Möller for their discussions, to W. R. Piggott of the Department of Scientific and Industrial Research, England, for his constant help and valuable contributions during the course of the work, to members of the Radio Research Station, Slough, England, for making simultaneous observations and to Dr. W. J. G. Beynon for assistance in preparing the final version of the paper.

REFERENCES

Appleton, E. V., Naismith, R., and Ingram, L. V., 1936, *Phil. Trans. Roy. Soc.* A, **236,** 191.

Beckmann, B., Menzel, W., and Vilbig, F., 1938 a, *Telegr.-Fernsprech-Funk-u. Fernsehtechnik*, **27,** 71; 1938 b, *Ibid.*, **27,** 245; 1939 a, *Ibid.*, **28,** 130; 1939 b, *Ibid.*, **28,** 223; 1940, *Ibid.*, **29,** 106; 1941, *Ibid.*, **30,** 43.

Böhm, O., 1929, *Telefunkenzeitung*, No. 53, 9.

Booker, H. G., and Wells, H. W., 1938, *Terr. Magn. Atmos. Elect.*, **43,** 249.

Dieminger, W., 1947, *Naturwissenschaften*, **1,** 29.

Dieminger, W., and Plendl, H., 1938, *Hochfrequenztech. u. Elecktroakust.*, **51,** 117.

Eckersley, T. L., 1929, *J. Instn. Elect. Engrs.*, **67,** 992; 1932, *Ibid.*, **71,** 405; 1937 a, *Nature, Lond.*, **140,** 846; 1937 b, *Gesammelte Vorträge der Hauptvers der Lilienthal Gesellschaft*, 307; 1937 c, *J. Instn. Elect. Engrs.*, **80,** 286.

Edwards, C. F., and Jansky, K. G., 1941, *Proc. Inst. Radio Engrs.*, **29,** 322.

Eyfrig, R., Goubau, G., Netzer, T., and Zenneck, J., 1938, *Hochfrequenztech. u. Elektroakust.*, **51,** 149.

Gerber, W., and Werthmüller, A., 1940, *Tech. Mitt.-Telegr.-Teleph.-Verw.*, **81,** 1.

Grosskopf, J., and Vogt, K., 1940, *Telegr.-Fernsprech-Funk-u. Fernsehtechnik*, **29,** 334.

Mögel, H., 1932, *Telefunkenzeitung*, No. 60, 29.

Netzer, T., 1940, *Hochfrequenztech. u. Elektroakust.*, **55,** 86.

Pinch Effect Oscillations in a High Current Toroidal Ring Discharge

By S. W. COUSINS and A. A. WARE
Department of Physics, Imperial College, London

Communicated by G. P. Thomson; *MS. received* 19*th September* 1950

ABSTRACT. A toroidal ring discharge with currents greater than 10^4 amperes has been studied with a rotating mirror camera. The photographs show a lateral oscillation of the discharge, which consists of an alternate contraction and expansion of the current filament. The frequency of the oscillations increases with decrease in pressure or increase in voltage and is inversely proportional to the square root of the atomic weight of the gas. The oscillations are thought to be due to plasma ion waves, which are excited by the large 'pinch' effect forces.

§ 1. INTRODUCTION

It is well known that in very high current gaseous discharges the self-magnetic field should cause a constriction of the discharge, but although this 'pinch' effect has been considered theoretically (Tonks 1937, 1939, Alfvén 1950) no experimental observation of the effect has been recorded.

Very recently, however, a form of toroidal ring discharge has been studied in which currents of the order of 10^4 amperes are produced (Ware 1951). The currents are excited by discharging a large condenser through a copper coating which is on the outside of the torus. Photographs of the discharge showed a marked constriction of the discharge in argon at pressures below 1 mm. Hg. With hydrogen the photographs showed uniform illumination across the tube with no apparent constriction.

This work has been continued, and a rotating mirror camera has been used to observe the temporal variation of the appearance of the discharge.

§ 2. DESCRIPTION OF APPARATUS

(i) *The Discharge Tube*

The apparatus used was similar to that described by Ware (1951). The diameter of the torus was 40 cm. and the cross section of the tubing 3 cm. To increase the inductance of the torus a laminated iron core was used, 80 cm. long and of cross section 10 cm. × 15 cm. This increased the inductance from 0·65 to 1·2 microhenries, and the fraction of the condenser voltage developed round the torus was 0·72.

The capacity used consisted of four 2·25 μF. condensers connected in parallel and capable of working up to a potential of 20 kilovolts, giving a maximum energy per flash of 1,800 joules. The resonant frequency of the circuit with no gaseous discharge was 42 kc/s.

An inspection window was made on the opposite side of the torus to the feed point by cutting a 1 mm. gap in the copper coating (Figure 1). To complete the electrical circuit a gauze cylinder was placed over the slit and strapped to the copper on either side.

Because of the disturbing effects caused by external magnetic fields the following precautions were taken to improve the screening of the gas: (*a*) The thickness of the copper coating was increased to about a millimetre. For a current

in the primary of 35,000 amperes the magnetic field under the copper was then reduced to about 1 gauss, the external field being about 1,000 gauss. (*b*) At the feed-point gap the gas was screened by means of a copper gauze cylinder which was under the copper coating (Figure 2). The gauze was insulated from the copper coating by means of a glass sleeve. To reduce the strong fields caused by the leads supplying the copper coating the leads were replaced by a pair of cones. The current was fed to the cones by two sets of leads, one on each side of the discharge tube. These cones produce only small fields in the gas and they act as an extra screen. (*c*) Small blocks of laminated iron linking the torus were placed near the feed point and the inspection window so that their demagnetizing fields helped to cancel the fields in the gas at these points.

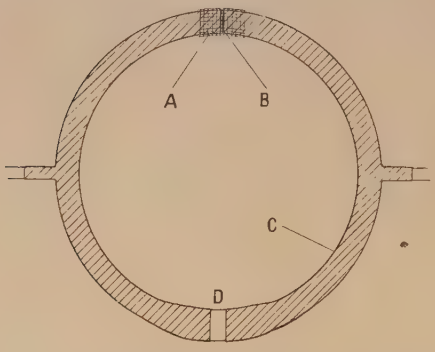


Figure 1. The torus. A—gauze cylinder, B—slit inspection window, C—copper coating, D—feed-point gap.

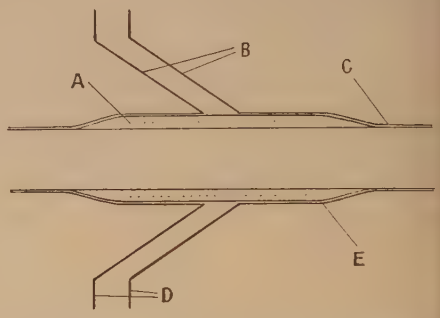


Figure 2. The feed point. A—gauze cylinder, B—gauze cones, C—glass torus, D—current leads, E—copper coating.

With these arrangements the magnetic field in the gas due to the primary current was reduced to about 10 gauss at the feed point and to about 20 gauss at the inspection window.

(ii) *The Rotating Mirror Camera*

The rotating mirror was a small polished 1-in. steel cube which was spun at 21,000 r.p.m. by a small induction motor.

An image of the inspection window slit was thrown on to a photographic plate via the rotating mirror; the rotating mirror photographs thus show the discharge at the slit traced out with respect to time. In order to fire the discharge when the mirror was in the correct position, a small spot of light was reflected from the mirror on to a photomultiplier cell; the voltage pulse produced was amplified and used to trigger the spark gap in the condenser discharge circuit.

Correlation of the photographs with the current oscillograms was obtained by throwing an image of the spark gap on to a slit above and collinear with the inspection window: the point on the photograph where the spark-gap light commences indicates the start of the condenser discharge.

(iii) *The Oscilloscope*

The method described by Ware was used to obtain a voltage proportional to dI_g/dt, where I_g is the gas current. The voltage was integrated by applying it to a resistance–capacity circuit of long time constant ($1{\cdot}5 \times 10^{-3}$ seconds), and the voltage across the capacity was amplified and applied to the Y plate of the oscilloscope; this voltage is proportional to the integral of the applied voltage.

§ 3. EXPERIMENTAL RESULTS

(i) *A General Description of the Photographs*

Examples of the rotating mirror photographs are shown in Plate I*. The horizontal black lines are caused by the wires of the gauze covering the inspection window. The gases studied in the torus were krypton, argon, oxygen, nitrogen, helium and hydrogen. The photographs fall into three groups, depending upon the atomic weights of the gases.

(a) Krypton and Argon

At all pressures in these gases an alternate contraction and expansion of the discharge is observed (Plate I, 1–4). Just after the breakdown the discharge, which is initially spread over the whole tube, is observed to contract into a narrow filament; the discharge does not stay 'pinched' but immediately expands again, and proceeds to oscillate in this way. These oscillations are not sinusoidal, contraction and expansion proceeding at fairly constant velocities. The discharge appears brightest at the outside edges, and in the contraction and expansion these bright edges trace out approximately straight lines. The discharge usually expands to the walls, but in some cases it stops expanding and begins to contract before reaching the walls: from the oscillograms it is found that this occurs only when very large currents are flowing in the gas. The discharge becomes brighter when it is contracted, and the brightness and sharpness of the 'pinch' increase with decrease in pressure. When the 'pinch' is very bright there is associated with it a bright line extending across the whole width of the tube.

Rotating mirror photographs were taken looking at both the top and the side of the discharge tube at the inspection window. For all except the lowest pressures the photographs were identical for the two views, showing that the oscillations were circularly symmetrical about the axis of the discharge tube. For pressures below about 0·1 mm. Hg, however, the photographs show transverse motions of the discharge superimposed upon the oscillations, so that 'pinches' are observed off the axis of the discharge tube. Also towards the end of these photographs diagonal streaks are observed which do not conform to any definite pattern.

(b) Oxygen and Nitrogen

The photographs for oxygen and nitrogen were almost identical at all pressures. Below 1·5 mm. Hg the photographs show oscillations similar to those in krypton and argon, except that they are of slightly higher frequency, but at the highest pressures the photographs are of the form shown in Plate I, 6 and 7. The gas appears to break down initially at the inside edge of the tube and the discharge then travels across the tube to the outside wall at a fairly constant velocity. It is then reflected back across the tube and finally fills the whole width. Plate I, 5, shows the discharge becoming narrow when it is near the walls of the discharge tube. This variation in width is caused by the circular cross section of the discharge tube.

(c) Helium and Hydrogen

In helium and hydrogen the oscillations are again observed, but they are of a higher frequency, which is near the limit of resolution of the photographs. Because of this, and the fact that less light is emitted by these gases, the

* For Plates see end of issue.

photographs are less distinct and the behaviour of the discharge is a little uncertain. At high pressures one or two oscillations can just be observed at the beginning of the discharge; these oscillations are symmetrical about the centre of the tube when viewed from the top, but the side views show a transverse motion superimposed on the oscillations (Plate I, 10 and 11). After these first few oscillations the photographs show no definite pattern.

(ii) *The 'Pinch' Effect Oscillations*

The frequency of these oscillations has been determined by measuring the periods between successive 'pinches'. Figure 3 shows the mean period plotted

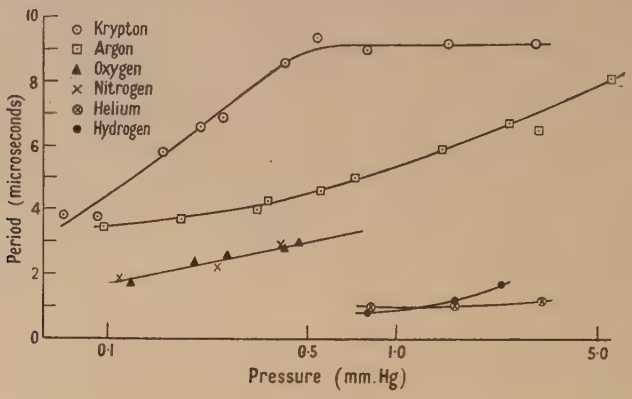


Figure 3.

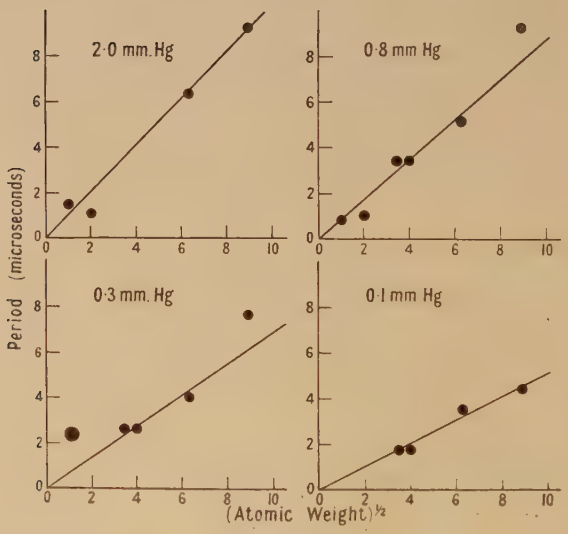


Figure 4.

against pressure for the gases studied, the condenser voltage being 15 kilovolts. The period is seen to decrease with decrease in pressure and is smaller the lighter the gas, the periods for nitrogen and oxygen being practically the same. In Figure 4 the periods for fixed pressures are plotted against the square root of the atomic weight of the gas. With the exception of krypton at medium pressures, the points fall approximately on straight lines passing through the origin, showing

that the period is directly proportional to the square root of the atomic weight. Most of the frequencies of the ' pinch ' oscillations are considerably greater than the frequency of the main gas-current oscillations, but at medium and high pressures in krypton the pinch oscillation period is exactly half the gas-current period and the 'pinches' occur approximately at the current peaks of every half-cycle. The fact that the period remains practically constant at this value for a considerable range of pressures in krypton (3·0 to 0·5 mm. Hg) suggests that the 'pinch' oscillations are pulled into phase with the gas current at these pressures.

The frequency of the 'pinch' oscillations increases slowly with increase in condenser voltage, for example with argon at 0·5 mm. Hg the periods for 7·5 and 15 kilovolts are 5·7 and 4·6 microseconds, respectively. Also for a given pressure, although the mean period is fairly constant, the individual readings vary considerably from one part of the trace to another. Comparison with the current oscillograms shows that the periods are shorter when the gas current is large.

Some rotating mirror photographs were also taken of the discharge at the feed-point gap, and the oscillations at this point were found to be in phase with those at the inspection window, the times of the 'pinches' at the two points coinciding to within about a twentieth of an oscillation. This is taken as sufficient evidence that the discharge filament contracts and expands as a whole throughout the torus.

(iii) *The Gas-Current Oscillograms*

Examples of the gas-current waveforms are shown in Plate II. Nos. 7–11 show corresponding oscillograms and rotating mirror photographs on the same time scale for comparison. These oscillograms are similar to those already described (Ware 1951) except that the frequency of the main oscillations is now about 55 kc/s. The peak amplitude of the gas current varies from 10^4 to 2×10^4 amperes.

The high-frequency oscillations described by Ware are again observed, and it has been found that these oscillations are closely connected with the 'pinch' oscillations. When comparing the oscillograms and the rotating mirror photographs it is found that when a ' pinch ' occurs there is a temporary decrease in the gas current (Plate II, 8 and 10). The effect varies with the sharpness of the 'pinch' and the amplitude of the gas current. It is most marked when the gas current is large and when the ' pinch ' is narrow and bright. In some cases the decrease is as much as 20% of the peak current, but in other cases where the 'pinch' is faint and the current small the effect is sometimes too small to be observed. The actual minimum of the gas current usually occurs slightly after the instant when the discharge is narrowest. These decreases in the gas current constitute the high-frequency oscillations on the oscillograms. They are best observed on the dI_g/dt waveforms, where the high-frequency variations have a larger amplitude relative to the main oscillations (Plate II, 11).

With krypton a very small decrease in the current is also observed when the discharge is expanded to the walls.

The effect is very small and is only observed on the dI_g/dt waveforms.

§ 4. CONCLUSIONS

The very large gas currents will cause very strong 'pinch' forces to occur in the gas. These forces will cause the current filament to contract until a

sufficiently large concentration gradient has been set up to counteract the 'pinch' forces. Because of the inertia of the positive ions it is possible for the discharge to oscillate about this equilibrium position, but such oscillations would be expected to be of an approximately sinusoidal nature, whereas rotating mirror photographs show the discharge contracting and expanding with a fairly constant velocity. This fact is more compatible with a wave motion in the gas, and it is suggested that the oscillations are due to a form of 'shock wave' in the plasma.

At the beginning of the discharge the 'pinch' forces will cause a compression of the plasma. This compression will be greatest at the outside edges of the tube where the 'pinch' forces are largest. (Also, if an appreciable skin effect occurs in the gas, as is likely if the ion concentrations are large, then the 'pinch' forces will only occur at the outside edges of the discharge.) The compression will be of considerable amplitude, and hence it will travel inwards as a 'shock wave' in the plasma. It will pass through the centre of the tube, travel out to the walls, and be reflected inwards again and so on. The bright parts of the discharge are identified with the wave front of the 'shock wave'. This brightness will be caused by the larger concentrations of ions in the wave front and, in addition, the 'shock wave' itself will give energy to the gas.

The 'shock waves' considered here are not large amplitude sound waves but large amplitude plasma ion waves. The theory of small amplitude plasma ion waves has been developed by J. J. and G. P. Thomson (1933) and the velocity of positive ion waves is given by

$$v = \left[\frac{k(T_e + T_p)}{M}\right]^{1/2} \qquad \ldots\ldots(1)$$

where T_e and T_p are the temperatures of the electrons and the positive ions respectively, k is Boltzmann's constant and M is the mass of a positive ion. The theory has been extended to the case of a plasma in a magnetic field by G. P. Thomson. It is found that the magnetic field causes a damping of the waves, but the velocity of the waves is still given to a good approximation by equation (1) provided $k_e H^2 e$ is small compared with Mp, where k_e is the electron mobility, H the magnetic field, e the electronic charge and p 2π the frequency of the waves. From analogy with normal shock waves, the velocity of ion waves of moderate amplitudes would be expected to be only a few per cent greater than the above velocity.

Assuming the velocity of the observed waves to be given by equation (1), the observed period between consecutive 'pinches' will be the time for the wave to travel the diameter of the discharge tube, and hence will be given by

$$\tau = \frac{d}{v} = d\left[\frac{M}{k(T_e + T_p)}\right]^{1/2} \qquad \ldots\ldots(2)$$

where d is the diameter of the discharge tube. τ should thus be proportional to $M^{1/2}$, which is in agreement with the experimental results. (The fact that straight lines were obtained in Figure 4 means that the added electron and positive ion temperatures were approximately the same in the different gases for a fixed pressure.) The slope of the straight line in Figure 4 should be equal to $d[k(T_e + T_p)]^{-1/2}$, and hence it is possible to calculate the sum of the electron and positive ion temperatures. These are found to vary from 15 ev. at 2 mm. Hg to 60 ev. at 0·1 mm. Hg.

In a normal shock wave the only forces acting on the molecules are those caused by collisions with other molecules. In an ion wave, however, the particles are charged and the wave can be affected by electric and magnetic fields. In the toroidal discharge the electrons in the 'shock wave' have a 'steady' component of velocity in the direction of the gas current. Because of this the magnetic field will exert a considerable force on them, pulling them towards the centre of the current filament. This force will be passed on to the positive ions by means of space charge forces. Hence, if a large gas current is flowing when the 'shock wave' is travelling outwards from the centre of the tube, these 'pinch' forces will stop the ions in the wave front and the 'shock wave' will be destroyed before it reaches the walls. (The wave is assumed to be a compression shock wave, so that ions in the wave front are travelling in the same direction as the wave.) The discharge will then begin to contract again and another 'shock wave' will be set up. This is considered to be the explanation of the cases where the 'shock wave' is observed to reverse direction before it reaches the walls. As expected, the effect is only observed when the gas current is large.

When the discharge is contracted to a narrow filament the self-inductance of the gas circuit will be considerably increased. Thus, for a ring with uniform current distribution, when the diameter of cross section is reduced from 3 to 1 cm. the self-inductance is increased from 0·65 to 0·93 microhenries. Hence the observed decrease in current when the discharge is contracted can easily be explained by the increase in inductance of the gas. The much smaller decrease in current which occurs in krypton when the discharge is expanded to the walls is probably due to an appreciable loss of ions at this point.

The patterns observed on the photographs for nitrogen and oxygen at high pressures are also thought to be due to these ion 'shock waves'. The difference in the patterns is caused by the initial breakdown occurring mainly at the inside edge of the tube. (The induced electric field causing the discharge is slightly greater at the inside edge of the tube, and this may be the cause of the gas breaking down more rapidly at this point.) The large current which is very quickly built up at this point will generate a high ion concentration there. This large concentration is equivalent to a high pressure, and hence an ion 'shock wave' will travel outwards from this point across the tube. It will be reflected by the outside wall, travel backwards across the tube, and so on. The observed velocities of this wave are the correct velocities expected for the 'shock wave' at these pressures.

The variations and movements of the discharge in hydrogen and helium are much more rapid than in the heavier gases. Because of this, and the fact that they emit less light, it has not been possible to make so accurate a study of the behaviour of the discharge in these gases. It is intended to use higher mirror speeds to obtain better results.

The characteristics of the rotating mirror photographs can thus be explained with the exception of the line of light associated with a bright 'pinch'. This line was at first thought to be due to reflection from the rear wall of the discharge tube of light from the bright central filament, but when the rear wall was blackened with lacquer there was only a small reduction in the intensity of the line. One possibility as to the origin of the line is fluorescence from the glass walls. This could be verified by combining a spectrograph with the rotating mirror camera to determine the nature of the light. Without such experimental

information it is not possible to form any definite conclusions as to the cause of the light.

ACKNOWLEDGMENT

The authors are indebted to Professor Sir George Thomson for his many helpful discussions and advice.

REFERENCES

Alfvén, H., 1950, *Cosmical Electrodynamics* (Oxford : Clarendon Press).
Thomson, J. J., and, Thomson, G. P., 1933, *Conduction of Electricity through Gases*, Vol. II (Cambridge : University Press), p. 355.
Tonks, L., 1937, *Trans. Electrochem. Soc.*, **72**, 167; 1939, *Phys. Rev.*, **56**, 360.
Ware, A. A., 1951, *Phil. Trans. Roy. Soc.* A, in the press.

Turbulence in Gaseous Conductors

By K. G. Emeléus
Department of Physics, Queen's University, Belfast

MS. received 1*st September* 1950

ABSTRACT. It is suggested that turbulent flow of electricity may occur under certain conditions in an ionized gas, even if the gas is not in a magnetic field. Examples are given of high and low frequency disturbances which may be of this character. The difficulties in the way of developing a quantitative theory, and the possibility of the coupling of electromagnetic and hydromechanical phenomena, are pointed out.

In an introduction to an account of work done during the war on electric discharges in magnetic fields which has been published recently, Bohm (1949) has stated that " a plasma in a magnetic field is always in a state of turbulent flow ". Direct experimental evidence for such turbulence has been obtained by Backus (1949). The possibility that turbulence can occur in gaseous conductors has also been visualized as a result of work done in this laboratory. It is thought that it may be of interest to indicate our lines of approach, which have been different from that of Bohm and Backus, and more general in that they do not obviously require the conductor to be placed in a magnetic field.

Two typical phenomena whose explanation has seemed difficult without the introduction of some new factor such as turbulence are the distribution in space of the ultra-high-frequency oscillations associated with primary electron beams, and the low-frequency (not greater than a few megacycles per second) noise generated by a plasma which is often, but not invariably, associated with the presence of unstable anode glows. The two frequency bands are those which have been shown by Tonks and Langmuir (1929), and by J. J. and G. P. Thomson (1933), to be characteristic of the coordinated electrostatic oscillation of small amplitude, of plasma electrons and ions respectively. They also appear in Bailey and Landecker's work (1950) on noise in the presence of a magnetic field. Both Bohm and Backus, and Massey and his collaborators (1949) in the detailed study of the discharges in magnetic fields have appreciated the significance of the two bands.

The peculiarity of the ultra-high-frequency oscillations is the abruptness with which they may appear at a certain distance from the cathode along the primary electron beam. This distance is several times the thickness of the space-charge sheath on the cathode. The existence of a species of discontinuity in the plasma was apparently first noticed by Druyvesteyn and Warmoltz (1937), who observed that there were sudden changes in the brightness of the discharge and in its spectrum. Later, Merrill and Webb (1939) found that ultra-high-frequency oscillations could be picked up by a probe to the anode side of a narrow sheet of plasma not more than 1 mm. thick, presumably at the optical discontinuity. Further experiments (Neill and Emeléus 1950) have confirmed and extended these observations, and shown that the behaviour of the primary beam is semi-quantitatively but not uniquely explained by Merrill and Webb's suggestion that it undergoes velocity modulation in a thin sheet of oscillating plasma at the junction. Such an explanation presupposes, however, the existence of the sheet, and it is difficult to see how any explanation can ignore the presence of it or something equivalent. It has been suggested (Emeléus 1949) that it may mark the abrupt build-up of an instability in the primary electron beam, very roughly analogous to what occurs when a stream of water from a chute has passed a short distance into a pool of water. In the discharge problem the interactions will certainly include an electrostatic term, and possibly (*infra*) an electromagnetic term.

The low-frequency phenomena are probably more complex than the ultra-high-frequency phenomena, in partial, but not exact analogy with positive ion plasma waves of small amplitude, which have group velocity features not shared by plasma electron oscillations of small amplitude (Tonks and Langmuir 1929). The one which has been most studied in this laboratory is the instability of anode glows. A particular form of glow may have only a limited range of existence conditions, and when the limit of these is reached, undergo a transition to another form. Relaxation oscillations may occur between the two forms for a short section of the average current–voltage characteristic curve for the tube (Armstrong 1942). Alternatively, a small glow may build up oscillations of approximately sinusoidal form and increasing amplitude, but before a steady state of oscillation has been reached, the system will collapse (Takamine *et al* 1933). In both cases, the plasma nearer the cathode becomes violently disturbed. In the former example, with the tubes studied, this is shown by noise and oscillations, in the latter by the appearance of moving striations. Neither disturbance of the plasma is capable of description as an ordinary positive-ion plasma wave, the amplitude of the disturbance being too large. The noise in the former case is also of higher level than is found in absence of an unstable anode glow or, presumably, other localized region of instability. The picture which has been tentatively formed of what is occurring is that the plasma disturbances, and possibly some of those of the anode spot itself, are turbulent. Strictly speaking, all that is shown definitely by the experiments is that disturbances of large amplitude occur; but it is an attractive extrapolation to suppose that, as in hydrodynamical phenomena, this may be accompanied by the production of circulation where it does not exist for disturbances of small amplitude. In other words, eddying is superposed on the drift current. The eddying could be, according to how it was produced, irregular, or, as in a Karman street, periodic. In the ultra-high-frequency example, if eddying occurs in the beam, it must also

occur in the motion of the 'ultimate' electrons in the plasma, at least in Merrill and Webb's oscillating sheet. Any eddying would give rise to local magnetic fields, and although the integrated field may be small or vanish outside the discharge, the local fields could serve as a means of coupling neighbouring eddies or neighbouring sections of the discharge, such as anode glow and main plasma. They could thus be of significance in two ways, viz. in building up eddies from small initial perturbations such as fluctuations in the magnetic field of the drift current, and in maintaining a statistical distribution of size and energy of the eddies once an assembly of them is established. Also, in a general sense, they correlate the phenomena which we have discussed with those considered by Bohm in which the magnetic field is applied from a source external to the discharge.

The quantitative development of these ideas, which will be essential before they can be thoroughly tested, will probably be even more elaborate than in hydrodynamics, because of the simultaneous existence in discharges of effectively two fluids, the electrons and positive ions. Some assistance in the development may be obtained by the use of dimensional methods, although these can only be applied when a fairly complete physical picture of what is taking place has been formed. Thus, if Reynolds' formula for the onset of turbulent flow in a tube of radius r cm. is adapted to the flow of electrons in a positive column, and approximate values are inserted for the free path of an electron for collisions with gas molecules, it follows that turbulence will set in when the drift current density is a multiple, probably between 25 and 50, of the random electron current density in the plasma divided by pr, where p is the pressure in mm. Hg. This supposes, however, that the effective density is that of the electron gas, and that the effective viscosity of the electrons is controlled by collisions with molecules, which, with other factors, requires closer examination.

It will also be necessary in each problem to examine whether or not the turbulent flow of electricity is associated with turbulent flow of gas. It is known (Langmuir 1923) that passage of a steady discharge can produce pressure differences in a gas, and it has been suggested (Emeléus and Gregg 1933) that in some discharges a coupling may occur between electrical and acoustical modes of vibration of the ionized gas. Turbulent motion of the electricity could be communicated to the gas in a similar way. Conversely, turbulent motion of a gas, such as can occur in sparks, will affect directly the ions and electrons which move with the gas. In the particular case of the violent disturbances which occur in a tube when ionization is almost complete (Langmuir and Mott-Smith 1924), the turbulent motions of the electricity and of the gas are one and the same thing. The combined electrical and mechanical aspects of turbulence may also be of importance in astrophysics, in regions with linear dimensions large compared with the electronic and molecular free paths. To give an example, Backus (1949) has shown that columns of plasma in a discharge in a magnetic field may detach themselves from the main plasma, and move away in some direction, a process which has a resemblance to the formation of solar prominences. It is also possible that collapse of space charges set up in a stellar atmosphere by differential rotation in the star's magnetic fields could contribute to the general turbulence of the atmosphere, and to the emission of radio noise (*cf.* also Ryle 1950).

It is scarcely necessary to emphasize that some of the ideas put forward here, and to a lesser extent those of Bohm and of Backus, are still speculative and partly

based on experiments far from completely studied. It is, however, a definite experimental fact that the instability of discharges of many forms is often associated with disturbances of large amplitude, and it seems worth investigating further if these are associated with rotation in the discharges.

REFERENCES

ARMSTRONG, E. B., 1942, *Thesis*, Belfast.

BACKUS, J., 1949, *The Characteristics of Electrical Discharges in Magnetic Fields*, edited by A. Guthrie and R. K. Wakerling (New York : McGraw-Hill Book Co.), chap. 11.

BAILEY, V. A., and LANDECKER, K., 1950, *Nature, Lond.*, **166**, 259.

BOHM, D., 1949, *The Characteristics of Electrical Discharges in Magnetic Fields*, edited by A. Guthrie and R. K. Wakerling (New York : McGraw-Hill Book Co.), chap. 1.

DRUYVESTEYN, M. J., and WARMOLTZ, N., 1937, *Physica*, **4**, 51.

EMELÉUS, K. G., 1949, Paper communicated to Conference on Gaseous Electronics, Pittsburgh, Nov. 3–5, 1949.

EMELÉUS, K. G., and GREGG, A. H., 1933, *Phil. Mag.*, **16**, 1079.

LANGMUIR, I., 1923, *J. Franklin Inst.*, **196**, 751.

LANGMUIR, I., and MOTT-SMITH, H. M., 1924, *Gen. Elect. Rev.*, **27**, 771.

MASSEY, H. S. W., *et al.*, 1949, *The Characteristics of Electrical Discharges in Magnetic Fields*, edited by A. Guthrie and R. K. Wakerling (New York : McGraw-Hill Book Co.), chap. 2 *seq.*

MERRILL, H. J., and WEBB, H. W., 1939, *Phys. Rev.*, **55**, 1191.

NEILL, T. R., and EMELÉUS, K. G., 1950, *Proc. Roy. Irish Acad.* A, **53**, 197.

RYLE, M., 1950, *Rep. Prog. Phys.*, **13** (London : Physical Society), p. 236.

TAKAMINE, T., SUGA, T., and YANAGIHARA, A., 1933, *Inst. Phys. and Chem. Res., Tokyo, Sci. Papers*, No. 403.

THOMSON, J. J., and THOMSON, G. P., 1933, *Conduction of Electricity Through Gases*, Vol. II, 3rd ed. (Cambridge : University Press), p. 353.

TONKS, L., and LANGMUIR, I., 1929, *Phys. Rev.*, **33**, 195.

CORRIGENDA

"The Variation with Temperature of the Piezoelectric Coefficients of Quartz", by A. C. LYNCH (*Proc. Phys. Soc.* B, 1950, **63**, 890).

Second line below Table 2, page 892 : 16_{60} *should read* 13_{60}.
Second line of Table 2: $0{\cdot}319_6$ *should read* $0{\cdot}0319_6$.

LETTERS TO THE EDITOR

Quantitative Measurement of Samples of Tritium

For measurements of certain nuclear disintegrations in this laboratory a quantitative determination of the amount of tritium present in the target gas became necessary. The apparatus described below was designed to do this. The use of Geiger counters to determine the rate of disintegration of tritium, and hence the concentration of tritium present, is governed by three considerations : (i) the short range of the β-particles from tritium involves putting the tritium inside the counter, (ii) the stopping power of the gas in the counter and the dimensions of the counter must be such that the wall effect is negligible, (iii) the sensitive volume of the counter must be determined.

To keep the wall effect small it was decided to use at least 10 cm. Hg pressure of argon in the counter and a diameter of $1\frac{7}{16}$ in. The possibility of an exchange with ordinary hydrogen makes the use of a hydrogenous quenching gas inadvisable when tritium is put inside a Geiger counter. With a mixture of 10 cm. argon and 1 cm. hydrogen it was found that spurious counts occurred up to 600 microseconds after the initial pulse; further, the pulse size in the absence of a quenching vapour was small, less than $\frac{1}{4}$ volt, and the Geiger region only extended about 50 volts; above this a continuous discharge started.

A quenching unit (for final form see Figure 1), with a quenching time of up to 1,000 microseconds, was constructed. The pulse, amplified by V_1 and V_2, triggered a pulse shaping circuit V_3 and V_4. The positive pulse of 100 volts was fed to the grid of V_5, biased to cut-off, the anode of which was connected to the counter; thus a negative pulse of up to

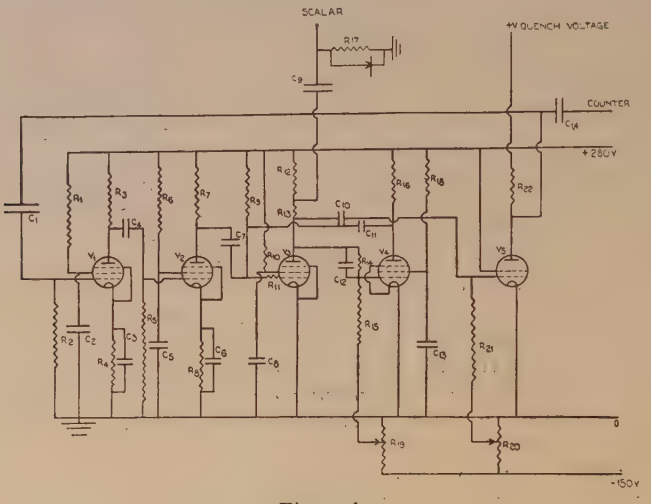


Figure 1.

V_1	6J7	C_5	0·5 μF.	C_{13}	0·5 μF.	R_7	10 kΩ	R_{15}	1M Ω
V_2	6J7	C_6	20 μF.	C_{14}	0·01 μF.	R_8	300 Ω	R_{16}	10 kΩ
V_3	VR65	C_7	15 pF.	R_1	68 kΩ	R_9	5 MΩ	R_{17}	10 kΩ
V_4	EF50	C_8	0·5 μF.	R_2	1M Ω	R_{10}	10 kΩ	R_{18}	10 kΩ
V_5	807	C_9	0·0003 μF.	R_3	10 kΩ	R_{11}	470 Ω	R_{19}	50 kΩ
C_1	0·0001 μF.	C_{10}	0·1 μF.	R_4	300 Ω	R_{12}	4·7 kΩ	R_{20}	50 kΩ
C_2	0·5 μF.	C_{11}	0·0005 μF.	R_5	1M Ω	R_{13}	10 kΩ	R_{21}	600 kΩ
C_3	20 μF.	C_{12}	50 pF.	R_6	68 kΩ	R_{14}	1M Ω	R_{22}	50 kΩ
C_4	0·0001 μF.								

2kv. could be applied to the counter wire soon after the initial pulse. It was found that spurious pulses were quenched provided the quenching pulse was greater than 250 volts, and no improvement was obtained from a larger pulse or by bringing the anode wire 100 volts negative to the cathode. A quenching pulse of about 280 volts was therefore used.

The sensitive volume of the counter was determined by using two counters identical in every respect except in the length of the central wire (Mann and Parkinson 1949). These were connected together and filled at the same time; the difference in the number of counts in the two counters was due to the difference in the lengths of the counters.

The counters (see Figure 2) were constructed of copper cathodes $1\frac{7}{16}$ in. diameter with a tungsten wire 0·006 in. in diameter as anodes. Thin stainless steel sleeves 0·02 in. diameter defined the lengths of the anodes, which were 5·03 cm. long in counter A and 10·12 cm. long

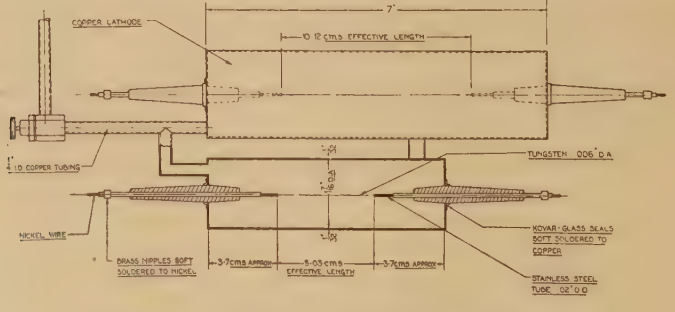


Figure 2. General arrangement of G–M counters.

in counter B. Owing to its relatively large diameter compared with its length, it was found that each counter had a slope to its plateau, 0·5% per volt and 0·25% per volt respectively; however, the difference in counts gave a plateau of zero slope within the experimental error of ±0·05% per volt. The slopes of the plateaux appear to be mainly due to end effects and would be reduced if counters longer in comparison with their diameters were used.

The measuring apparatus was tested for linearity with various concentrations of tritium; the results corrected for dead time and normal background are shown in Figure 3.

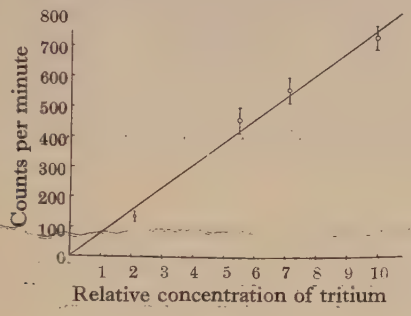


Figure 3.

Discussion.

The quenching circuit used was the simplest that gave satisfactory results. To prevent blocking at high counting rates it is necessary to keep the counter voltage below the point where a continuous discharge sets in before the quenching circuit has recovered from the previous pulse (Collinge 1950). The working plateaux were thus limited to about 50 volts, the lower limit being set by the fact that the amplification was kept low so that excessive screening from outside interference was not required. Using copper cathodes, it was possible to remove the tritium easily, and the normal background counts were achieved after a few minutes pumping. The counter-potential was about 1,000 volts. The long quenching pulse restricts the counting rate to three thousand a minute.

With the comparatively simple apparatus described above, measurements of the absolute concentration of tritium can be made.

Clarendon Laboratory, J. H. Frisby.
Oxford. D. Roaf.
28th November 1950.

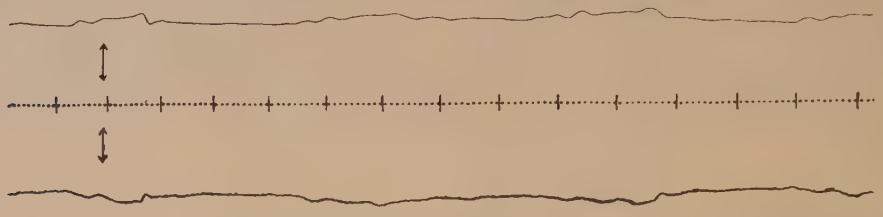
Collinge, B., 1950, *Proc. Phys. Soc.* B, **63**, 665.
Mann, W. B. and Parkinson, G. B., 1949, *Rev. Sci. Instrum.*, **20**, 41.

Binocular Eye Movements at High Convergence

The photoelectric corneal reflex method has been applied to the measurement of binocular eye movements of subjects in the sitting position (Lord 1951): in principle, the method is similar to that used in investigations on supine subjects (Lord 1948). When the subject looks binocularly at the fixation target, each cornea is irradiated by an ultra-violet beam (λ 3650) directed towards the blind spot, so that the subject is practically unaware of the ultra-violet beams. The fixation target is situated on the visual axis of the subject's right eye when this eye is looking straight ahead; in the present instance the target was 18 in. from the subject. In order that the left-eye ultra-violet beam may fall on the blind spot for other target distances, provision has been made for adjusting the direction of this beam. The movements of the ultra-violet beam reflected at each cornea are recorded as described by Lord (1948, 1951); so far, the investigations have been confined to simultaneous recordings of the movements of both eyes in one direction at a time. To position the subject,

adjustments are made for the right and left eyes in turn until the correct positions of both eyes are secured simultaneously: the horizontal separation of the right and left eye optical systems is adjustable and the dental impression mouthpiece, on which the subject bites, can be rotated about horizontal and vertical axes passing through the centre of rotation of the subject's right eye; the mouthpiece is capable of the conventional three mutually perpendicular translations.

Head-movement records obtained while the subject is fixating are very similar to those previously obtained with the subject upright (Lord 1950, 1951). The movements are periodic, with average maximum value 0·03 mm. in the side-to-side direction.



Up-and-down eye records for subject R.S.L. for binocular fixation of a bright disc, angular diameter 1 min., distant 18 in. from the subject. Top, left eye; bottom, right eye; towards time line corresponds to upward movement. Time dots, 50 per second. Height of arrowed line represents a rotation of 20 minutes, assuming the separation of the centre of curvature of the cornea and the centre of rotation of the eye is 5·3 mm.

There are several points of resemblance between the binocular records for high convergence and those (Lord 1951) obtained with a target distance of 5½ ft. The head-movement pattern can be picked out in almost all records. Moreover, periodic up-and-down head movements deduced from left and right eye records are again out of phase; this has been attributed to head movement about a horizontal axis perpendicular to the front of the subject. It again appears that, within the limitations of the apparatus, the eyes move at the same time and by the same amount; further, the movements are in the same direction. As in the case of small convergence, flicks starting in the downward direction are almost completely absent. For all subjects, flicks starting to the right are about twice as frequent as those starting to the left; this is also the case at low convergence, except for subject W.D.W., for whom the flicks occurred with about equal frequency. The value of the average interval between flicks is 0·86, 0·97, 2·44 and *c.* 10 seconds for R.W.G.H., W.D.W., R.S.L. and L.C.T. respectively. The value for R.W.G.H. is again unchanged: the invariance (Lord 1950, 1951) of the average flick frequency of this subject is very striking. The average flick interval of L.C.T. is also unaffected by the change in convergence; that for W.D.W. is slightly increased, but is smaller than that found for monocular fixation for this subject in the upright position (Lord 1950). In the case of R.S.L. (see Figure), increase in convergence has a most marked effect on the records: the average interval at small convergence is about 10 seconds. Moreover, the subject is aware of the phenomenon and attributes it to the cultivated habit of periodically relaxing his accommodation while performing close work. The duration and magnitude of the flicks (Lord and Wright 1948, Lord 1950, 1951), are again 0·02–0·03 second and from 2–25 minutes of arc respectively.

The author is grateful to Dr. W. D. Wright for many helpful criticisms and for continuing to act as subject. She much appreciates the sustained interest of the other subjects, Dr. L. C. Thomson, Dr. R. S. Longhurst and Mr. R. W. G. Hunt, with all of whom she has had many interesting discussions. The continued financial support of the Medical Research Council is gratefully acknowledged.

Chelsea Polytechnic, London S.W. 3, and
Imperial College, London S.W. 7.
6th December 1950.

Mary P. Lord.

Lord, M. P., 1948, *Proc. Phys. Soc.*, **61**, 489; 1950, *Nature*, **166**, 349; 1951, *Brit. J. Ophthal.*, in the press.
Lord, M. P., and Wright, W. D., 1948, *Nature, Lond.*, **162**, 25.

REVIEWS OF BOOKS

Optique Physiologique, Tome Deuxième: Lumière et Couleurs, par Y. Le Grand. 2nd Edition. Pp. 490. (Paris: Éditions de la *Revue d'Optique*, 1950.) 1,600 fr.

A number of books on vision and colour have appeared during the last few years and, while comparisons are invidious, the present volume by Le Grand must be regarded as quite outstanding. In assessing its merits it should, however, be remembered that the general title of the three volumes (of which this is the second) is *Physiological Optics*. The reader must not, therefore, expect to find the more technical aspects of colorimetry described in great detail; thus spectrophotometry is virtually omitted, while colour atlases are dealt with in a paragraph or so. It follows that the book is very different in scope and outlook from, say, Evans' *Color*, neither does it attempt to compete in lavishness of production with that excellent book.

Yet it would be difficult to imagine a more thoughtful study of vision than Le Grand has given us. His own original researches are by no means negligible, but he has, in addition, a remarkable insight into the significance of research work other than his own. In the section dealing with colour mixture data, for example, this insight has evidently been gained from detailed calculations and transformations which he has himself made on the data; as a consequence, the reader is unlikely to encounter a discussion on this subject of greater clarity, while anyone who has any lurking inhibitions about the significance of negative quantities in colour equations could hardly do better than go to Le Grand to have them out.

In the chapters dealing with discrimination we find a useful discussion of the methods by which the difference lumen can be measured, and the final chapter of the book consists of an interesting survey and commentary on the theories of thresholds which have been proposed from time to time. Perhaps Le Grand tends to regard discrimination too much as a function of the physiological processes and to discount or disregard the psychological aspects of the problem. But his amusing reference on p. 276 to the troubles which Fechner, " l'honnête professeur de physique à Leipzig ", brought down on his head by his celebrated deduction that the sensation is proportional to the logarithm of the stimulus, suggests that Fechner's experience may have deterred Le Grand from advancing too far into this psychological territory.

To list the other subjects considered in the book—light sources, photometric quantities, luminosity curves, temporal and spatial effects, defective colour vision, the physiology of the retina, colour vision theories, and so on—must obviously do less than justice to the contribution which Le Grand makes to these matters. One of the most valuable and intriguing features is the delicate criticism—almost an aside—which is from time to time thrown out: thus on p. 78 he voices his dislike of the term ' illuminance ', proposed in the United States of America, on account of its confusion with luminance, a dislike shared by many in this country; on p. 93 he throws reasonable doubt on the data by Ludvigh and MacCarthy of the spectral transmission values of the ocular media; on p. 170 he voices the opinion (strongly held also by the Terminology Committee of the Colour Group of the Physical Society) that it is undesirable to try to limit the use of the word ' colour ' to only one of its possible meanings; he occasionally offers a mild rebuke to those workers who let their enthusiasm for mathematics lead them into complicated theories for which there is little or no experimental justification; and if, on p. 322, Le Grand sees fit to doubt the merits of a unit of subjective brightness, the ' brill ', which was once proposed by the reviewer, then such is the reviewer's regard for Le Grand's opinion that he is almost (but not quite) persuaded that Le Grand must be right!

This is undoubtedly a book which all workers on vision will read with much profit.

W. D. WRIGHT.

CONTENTS FOR SECTION A

ABSTRACTS FOR SECTION A

The Theory of the Propagation in Liquid Helium II of 'Temperature-Waves' of Finite Amplitude, by H. N. V. Temperley.

ABSTRACT. The two-fluid model of liquid helium II is applied to the problem of predicting the behaviour of waves and pulses of second sound of finite amplitude. It is found that equations quite analogous to the Riemann and Rankine–Hugoniot equations for ordinary sound waves and shock fronts can be established. These equations predict the 'overtaking effect' with the establishment of a 'temperature-shock' for pulses of second sound of finite amplitude. The temperature-shocks should be of 'front-edge' type (propagation of an abrupt rise in temperature) if the liquid helium is at a temperature below 1·96° K., but should become 'back-edge', corresponding to an abrupt fall in temperature, between 1·96° K. and the λ-point. There is experimental evidence in support of both these predictions. It is shown that, just as for ordinary shock fronts, the irreversible effects due to the passage of a temperature shock should be proportional to the cube of the height of the shock.

Second Sound in Liquid Helium II, by D. V. Osborne.

ABSTRACT. A pulse technique for the propagation of second sound is described and the expected shape of the received pulse is discussed. In this connection, experimental evidence has been obtained for a temperature discontinuity at the interface between liquid helium II and a solid medium at second sound frequencies. The pulse method has been used to study the attenuation of second sound with distance and the formation of shock waves. The observations on the latter are shown to be in agreement with the predictions of the equations of motion of liquid helium II if second order terms are not neglected.

The Magnetic Susceptibility of Copper Sulphate, by R. J. Benzie and A. H. Cooke.

ABSTRACT. Measurements of the magnetic susceptibility of copper sulphate pentahydrate have been made at temperatures between 20° K. and 1° K. on a single crystal specimen. It is found that the susceptibilities measured in different directions with respect to the crystallographic axes follow a Curie–Weiss law, with different Curie constants but a fixed Weiss constant of −0·6°. The discrepancy between this result and that of Krishnan and Mookherji is attributed to temperature-independent paramagnetism.

Theory of the Superconducting State: II—*Magnetic Properties at the Absolute Zero of Temperature*, by H. Fröhlich.

ABSTRACT. It has previously been shown by the author that the interaction between free electrons and the lattice vibrational field (if strong enough) leads at the absolute zero to the formation of a new state. The London equations are now derived in the sense conjectured by F. London.

The Second Order Photoelectric Effect at a Metal Surface, by R. E. B. Makinson and M. J. Buckingham.

ABSTRACT. An expression is derived for the photoelectric current excited at the surface of a metal by radiation of frequency below the threshold frequency of the first order photoelectric effect. A physical interpretation of the formal results is given and the possibility of detecting this current experimentally is briefly discussed.

On the Theory of Plastic Deformation, by Doris Kuhlmann.

ABSTRACT. A model for plastic deformation in soft metals, i.e. metals which are not age-hardened or severely cold-worked, is given. It is based on the assumption that in real crystals there is always a number of regions present which, by an unknown process, act as sources of dislocations. Furthermore, it is assumed that in soft metals dislocations are very mobile except in a limited number of regions which act as obstacles. The model is applied to explain a number of facts connected with plastic deformation.

Factors Controlling Dislocation Widths, by A. J. Foreman, M. A. Jaswon and J. K. Wood.

ABSTRACT. The work of Peierls and Nabarro is extended to a family of edge dislocations of greater widths. The weaker the shearing forces between adjacent atoms in the slip plane, the wider is the dislocation. The external shear stress required to move the dislocations is extremely sensitive to the width, becoming vanishingly small at widths of the order of three atomic spacings. The theory is applied to bubble raft dislocations, and satisfactory agreement is found with experiment.

Further Studies—Experimental and Theoretical—of a Wide-Angle β-Spectrometer which uses a Prolate Spheroidal Magnetic Field, by T. H. Braid and H. O. W. Richardson.

ABSTRACT. A β-spectrometer is described which focuses a conical sheaf of rays emitted at about 80° to the axis of a prolate spheroidal magnetic field. In agreement with a previous theoretical prediction, the trajectories cross over at two sites in the rotating meridional plane. The field behaves in some ways like a magnetic mirror and can be arranged to form an aberration-free image of a point on the axis of symmetry. It cannot, however, give a clear image of an area because the Abbe sine condition is not fulfilled.

The wide-angle focusing action is compared with that recently found by Slätis and Siegbahn, for which a different optical analogue is suggested, based on the isolation of a meridional focal line between two refractions.

The adjustment of the spectrometer and measurements of its resolving power and solid angle of collection are described, an example being 1·4% of 4π with a line-width of 1/113 at half-height. By working at less than unit magnification it is possible to focus rays emitted at less than 80° to the axis.

Observations of Slow Mesons and Nuclear Disintegrations in Photographic Plates exposed under Carbon Absorbers, by J. C. Barton, E. P. George and A. C. Jason.

ABSTRACT. The exposure was made at the Jungfraujoch, altitude 3,457 m., using Ilford type C2 nuclear research emulsions. The frequency of stars under thicknesses of carbon up to 220 gm/cm² was measured, the experimental value of the absorption length of the radiation causing them being 166 ± 8 gm/cm². The proportion of the stars due to fast π-mesons has been evaluated and, on subtracting this proportion, the absorption length of the radiation responsible for the remainder is found to be 143 ± 10 gm/cm². The energy distribution of the stars having from 3 to 20 tracks is found to be of the form $dE/E^{2\cdot6}$, with no apparent discontinuity. The frequencies of slow π- and μ-mesons at various depths is reported. The π-mesons exhibit a transition effect which is interpreted in terms of their production in stars occurring in the carbon absorber. The ratio of negative to positive π-mesons is found to be 3·1 to 1.

Disintegrations produced by the Nuclear Capture of Slow Negative μ-Mesons, by E. P. George and J. Evans.

ABSTRACT. The 'prong distribution' for the disintegrations produced by the σ-mesons of the cosmic radiation in photographic emulsions exposed below ground is found to be different from that observed in emulsions exposed above ground. The underground distribution contains relatively more 'stars' consisting of one prong. The observations are consistent with the hypothesis that most of the σ-mesons below ground are in fact negative μ-mesons and not negative π-mesons. It is concluded that $8\cdot7 \pm 1\cdot7\%$ of the negative μ-mesons captured in the emulsion lead to disintegrations accompanied by the emission of charged particles. The average number of prongs per disintegration produced by the capture of negative μ-mesons is $0\cdot10 \pm 0\cdot02$. The mean number of prongs and the associated prong distribution suggest that a nuclear excitation of about 15 Mev. is produced by the capture of a μ-meson.

Some Gamma-Rays from Light Elements under Proton Bombardment, by J. H. Carver and D. H. Wilkinson.

ABSTRACT. The energy of the photo-protons from deuterium, measured in an ionization chamber, has been used to determine the energy of gamma-rays produced by the proton bombardment of several light elements. The results are expressed in the following Table :

Isotope	Target thickness (kev.)	Proton energy (kev.)	Gamma-ray energy (Mev.)
^{9}Be	2	1000	$7\cdot39 \pm 0\cdot15$
^{11}B	200	950	$12\cdot50 \pm 0\cdot21$
^{13}C	20	570	$8\cdot14 \pm 0\cdot08$ and $5\cdot81 \pm 0\cdot25$
^{19}F	thick	740	$12\cdot09 \pm 0\cdot28$
^{7}Li	thick	500	$14\cdot4 \pm 0\cdot4$
^{15}N	unknown	900	$4\cdot45 \pm 0\cdot04$

The absence of certain well-known lines from the above table merely means that their energy was not determined in this investigation.

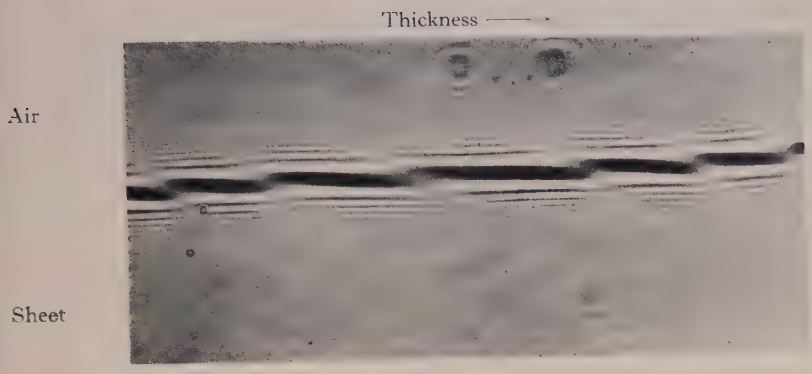


Figure 3. (×60)

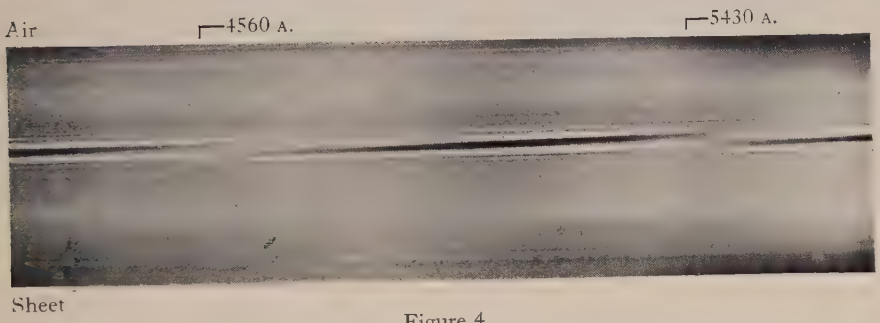


Figure 4.

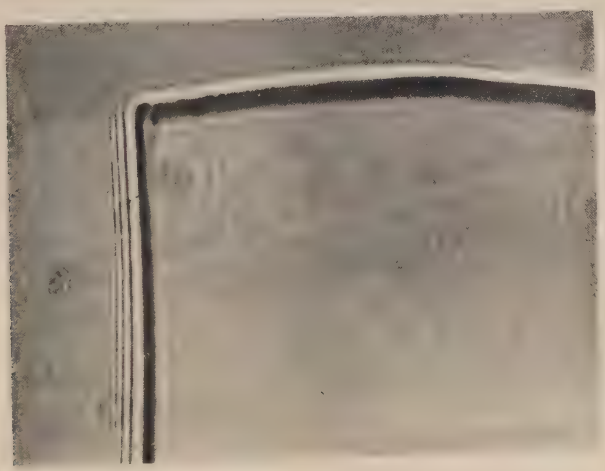


Figure 6. (×120)

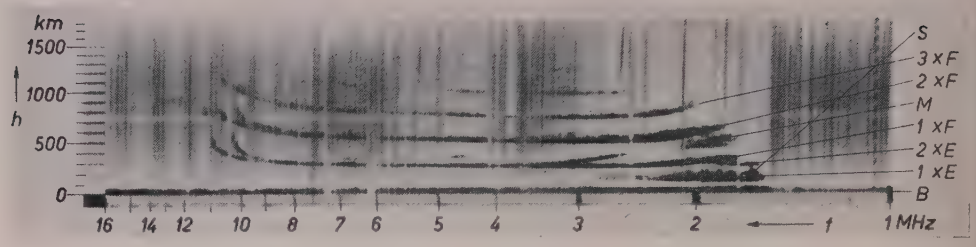


Figure 2. Multi-frequency record with E scatter at sunset. 8th March 1948. 1800 Middle European time (M.E.T.).

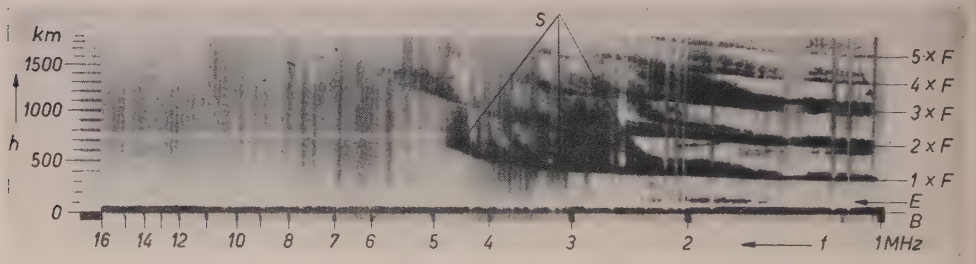


Figure 3. Multi-frequency record with $1 \times F$ scatter at night-time. 14th March 1948. 0430 M.E.T.

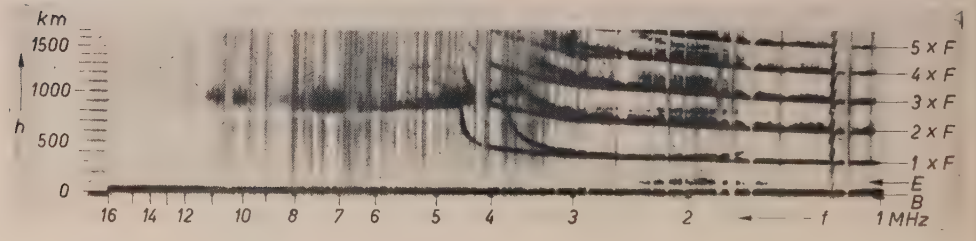


Figure 4. Multi-frequency record with G scatter. 28th February 1948. 0400 M.E.T.

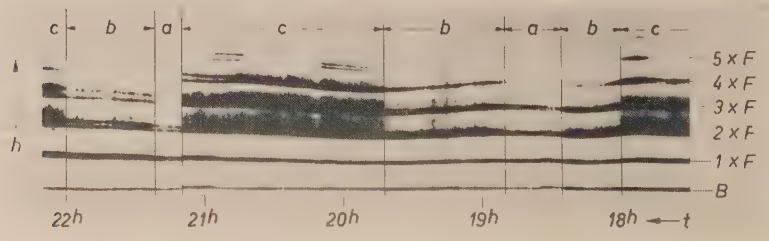


Figure 5. Record of a fixed frequency (7 Mc/s., 50 kw.), taken with different sensitivities of the receiver. (*c*) High, (*b*) medium, (*a*) low. 27th August 1946. 2000 M.E.T.

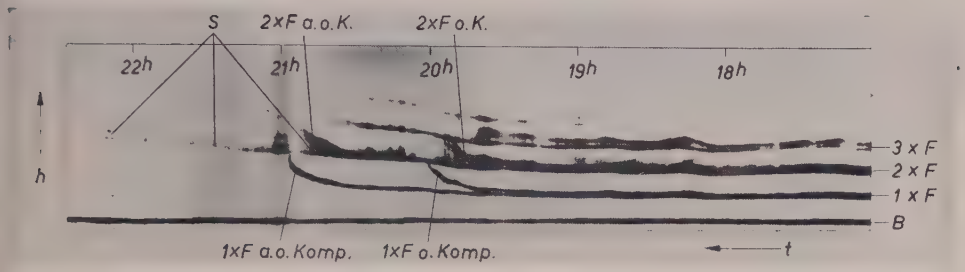


Figure 6. Record of a fixed frequency (7 Mc/s., 50 kw.) when the used frequency approaches the critical one; S–scattered echoes. 6th September 1946. 1730–2230 M.E.T.

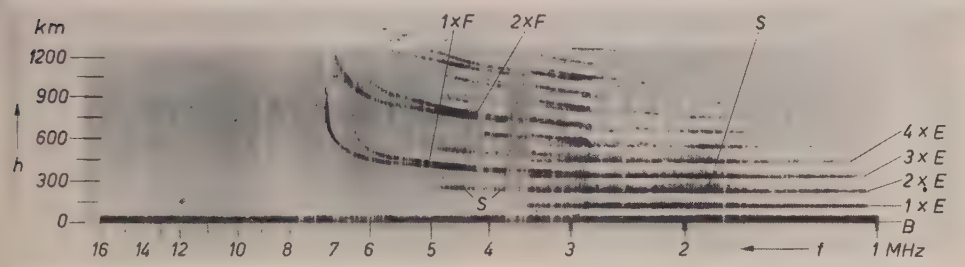


Figure 11. Multi-frequency record with 2×E scatter extending to frequencies higher than the critical frequency of 1×E. 14th May 1948. 0130 M.E.T.

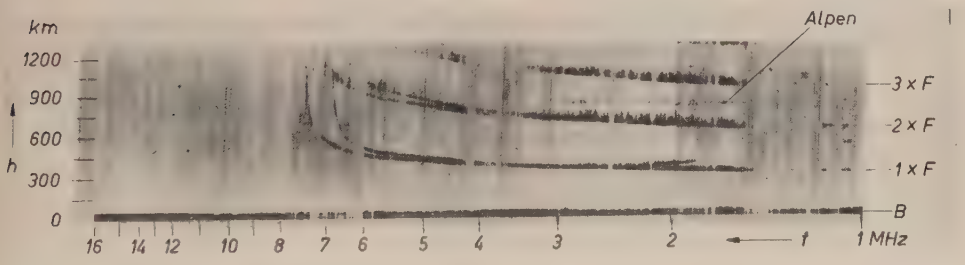
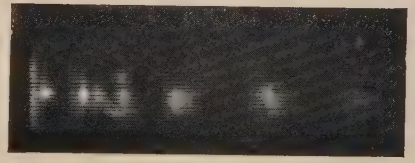


Figure 12. Multi-frequency record with a marked maximum of intensity of scatter corresponding to a ground distance of the radiation of about 500 km. 16th May 1948. 2330 M.E.T.

KRYPTON

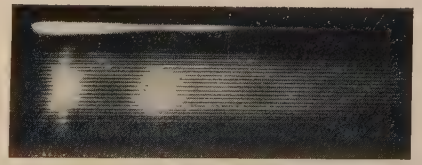
1. 0·6 mm. Hg, 20 kv.

2. 0.24 mm. Hg, 20 kv.



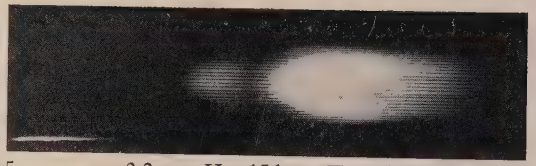
3. 0·05 mm. Hg, 20 kv.

ARGON

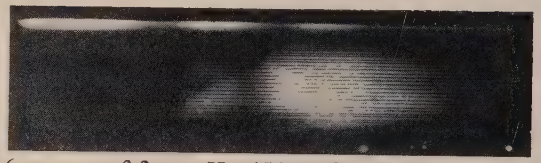


4. 0·55 mm. Hg, 15 kv.

OXYGEN

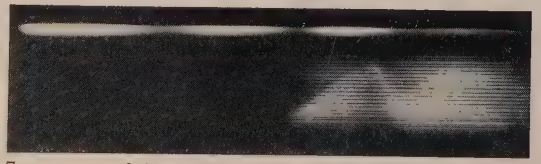


5. 3·2 mm Hg, 15 kv. Top view.

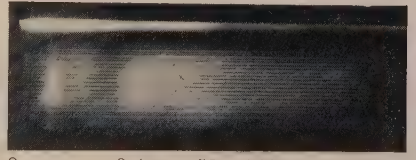


6. 3·2 mm. Hg, 15 kv. Side view.

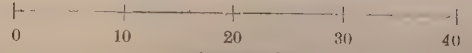
NITROGEN



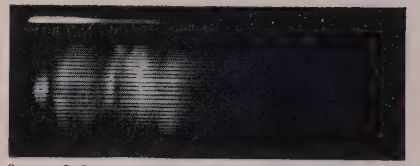
7. 3·1 mm. Hg, 15 kv. Side view.



8. 0·4 mm. Hg, 15 kv.

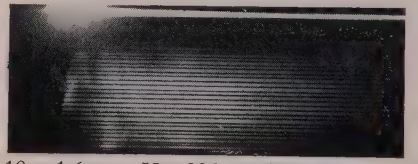


HELIUM



9. 0·8 mm. Hg, 20 kv. Top view.

HYDROGEN



10. 1·6 mm. Hg, 20 kv. Side view.

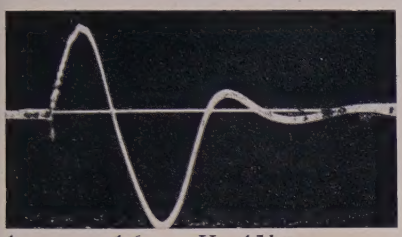
11. 0·8 mm. Hg, 20 kv. Side view.

12. 0·8 mm. Hg, 20 kv. Top view.

PLATE I

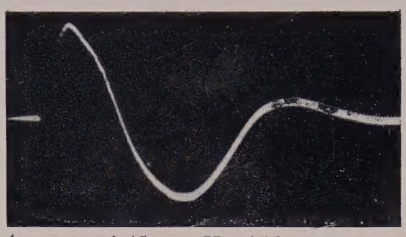
HYDROGEN.

ARGON

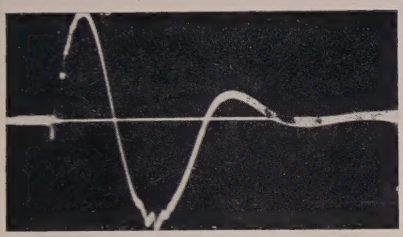


1. 1·6 mm. Hg, 15 kv.

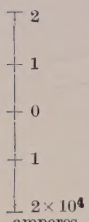
Gas current waveforms

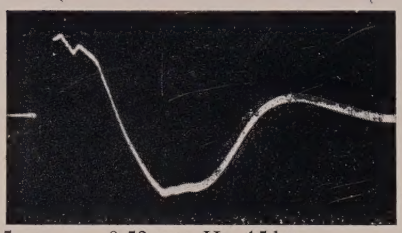


4. 1·45 mm. Hg, 15 kv.

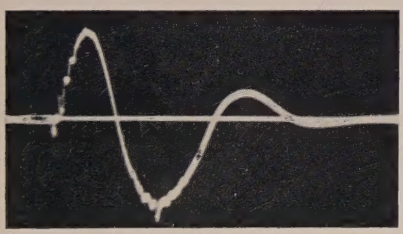


2. 0·4 mm. Hg, 15 kv.

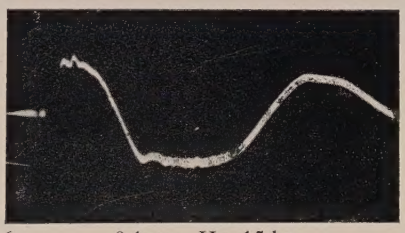




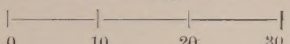
5. 0·52 mm. Hg, 15 kv.

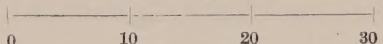


3. 0·22 mm. Hg, 15 kv.

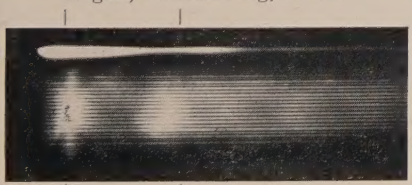


6. 0·1 mm. Hg, 15 kv.

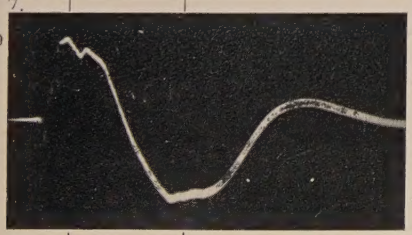




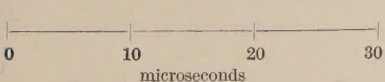
Argon, 0·52 mm. Hg, 15 kv.



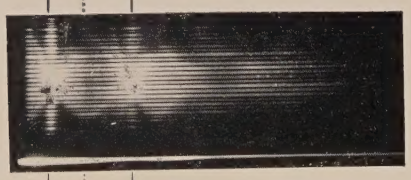
7.



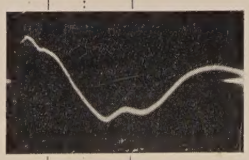
8. Gas current waveform.



Krypton, 0·2 mm. Hg, 15 kv.

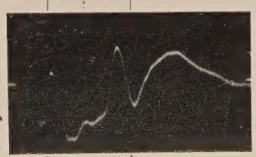


9.



Gas current waveform.

10.



$\frac{dI_g}{dt}$ waveform

11.

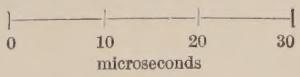


PLATE II.

THE PHYSICAL SOCIETY

MEMBERSHIP

Membership of the Society is open to all who are interested in Physics:

FELLOWSHIP. A candidate for election to Fellowship must as a rule be recommended by three Fellows, to two of whom he is known personally. Fellows may attend all meetings of the Society, are entitled to receive Publications 1 (either Section A or Section B), 4 and 5 below, and may obtain the other publications at much reduced rates.

STUDENT MEMBERSHIP. A candidate for election to Student Membership must be between 18 and 26 years of age and must be recommended from personal knowledge by a Fellow. Student Members may attend all meetings of the Society, are entitled to receive Publications 1 (either Section A or Section B) and 4, and may obtain the other publications at much reduced rates.

Books and periodicals may be read in the Society's Library, and a limited number of books may be borrowed by Fellows and Student Members on application to the Honorary Librarian.

Fellows and Student Members may become members of the *Colour Group*, the *Optical Group*, the *Low Temperature Group* and the *Acoustics Group* (specialist Groups formed in the Society) without payment of additional annual subscription.

PUBLICATIONS

1. *The Proceedings of the Physical Society*, published monthly in two Sections, contains original papers, lectures by specialists, reports of discussions and of demonstrations, and book reviews. Section A contains papers mainly on atomic and sub-atomic subjects; Section B contains papers on macroscopic physics.

2. *Reports on Progress in Physics*, published annually, is a comprehensive review by qualified physicists.

3. *The Handbook of the Physical Society's Annual Exhibition of Scientific Instruments and Apparatus*. This Exhibition is recognized as the most important function of its kind, and the Handbook is a valuable book of reference.

4. *The Bulletin*, issued at frequent intervals during the session, informs members of programmes of future meetings and of the business of the Society generally.

5. *Physics Abstracts* (*Science Abstracts A*), published monthly in association with the Institution of Electrical Engineers, covers the whole field of contemporary physical research.

6. *Electrical Engineering Abstracts* (*Science Abstracts B*), published monthly in association with the Institution of Electrical Engineers, covers the whole field of contemporary research in electrical engineering.

7. *Special Publications*, critical monographs and reports on special subjects prepared by experts or committees, are issued from time to time.

MEETINGS

At approximately monthly intervals throughout each annual session, meetings are held for the reading and discussion of papers, for lectures, and for experimental demonstrations. Special lectures include: the *Guthrie Lecture*, in memory of the founder of the Society, given annually by a physicist of international reputation; the *Thomas Young Oration*, given biennially on an optical subject; the *Charles Chree Address*, given biennially on Geomagnetism, Atmospheric Electricity, or a cognate subject; and the biennial *Rutherford Memorial Lecture*. Meetings are generally held each year at provincial centres, and from time to time meetings are arranged jointly with other Societies for the discussion of subjects of common interest.

Each of the four specialist Groups holds about five meetings in each session.

SUBSCRIPTIONS

Fellows pay an Entrance Fee of £1 1*s*. and an Annual Subscription of £3 3*s*. Student Members pay only an Annual Subscription of 15*s*. Second Section of *Proceedings* 30*s*. No entrance fee is payable by a Student Member on transfer to Fellowship.

Further information may be obtained from the Secretary-Editor at the Offices of the Society:

1 LOWTHER GARDENS, PRINCE CONSORT ROAD, LONDON S.W. 7
Telephone: KENsington 0048, 0049

Printed by TAYLOR AND FRANCIS, LTD., Red Lion Court, Fleet Street, London E.C.4.